CANADIAN THESES ON MICROFICHE

I.S.B.N.

THESES CANADIENNES SUR MICROFICHE

National Library of Canada
Collections Development Branch

Canadian Theses on
Microfiche Service

Ottawa, Canada
K1A 0N4

Bibliothèque nationale du Canada
Direction du développement des collections

Service des thèses canadiennes
sur microfiche

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us a poor photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de mauvaise qualité.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE

NL-339 (r. 82/08)

Canada

0-315-08923-7

National Library of Canada | Bibliothèque nationale du Canada

34 Canadian Theses Division | Division des thèses canadiennes

Ottawa, Canada
K1A 0N4

56840

PERMISSION TO MICROFILM — AUTORISATION DE MICROFILMER

• Please print or type — Écrire en lettres moulées ou dactylographier

Full Name of Author — Nom complet de l'auteur

EVANS, DAVID HUGH

Date of Birth — Date de naissance	Country of Birth — Lieu de naissance
26 JULY 1956	MALAWI

Permanent Address — Résidence fixe

15907 -78 AVE
EDMONTON, ALTA
CANADA T5R 3E2

Title of Thesis — Titre de la thèse

UNUSUAL ASPECTS OF POLYNUCLEOTIDE STRUCTURE

University — Université

ALBERTA

Degree for which thesis was presented — Grade pour lequel cette thèse fut présentée

Ph. D.

Year this degree conferred — Année d'obtention de ce grade	Name of Supervisor — Nom du directeur de thèse
1982	DR. A. R. MORGAN

Permission is hereby granted to the NATIONAL LIBRARY OF CANADA to microfilm this thesis and to lend or sell copies of the film.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

L'autorisation est, par la présente, accordée à la BIBLIOTHÈQUE NATIONALE DU CANADA de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans l'autorisation écrite de l'auteur.

Date	Signature
MARCH 11, 1982	D. Evans

NL-91 (4/77)

THE UNIVERSITY OF ALBERTA

Unusual Aspects of Polynucleotide Structure

by

© David H. Evans

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

of Doctor of Philosophy

Biochemistry

University of Alberta

Spring, 1982

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR David H. Evans

TITLE OF THESIS Unusual Aspects of Polynucleotide Structure

DEGREE FOR WHICH THESIS WAS PRESENTED Doctor of Philosophy

YEAR THIS DEGREE GRANTED Spring 1982

Permission is hereby granted to THE UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

(SIGNED) D. Evans

PERMANENT ADDRESS:

15907-78 Ave

Edmonton Alberta

T5R 3E2

DATED February 26 1982

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled Unusual Aspects of Polynucleotide Structure submitted by David H. Evans in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

A. R. Morgan
Supervisor

Joel H. Weiner

Linda J. Reha-Krantz

Cyril M. Kay

Stuart M. Linn
External Examiner

Date March 11, 1982

Abstract

A number of aspects of polynucleotide structure are investigated using synthetic repeating sequence polynucleotides. RNA's containing a repeating purine base-sequence form four-stranded structures analogous to those formed by polypurine DNA's and by poly rG. Synthetic RNA's are difficult to prepare in quantity and the reason appears to be formation of inert DNA·RNA·RNA polymerase complexes during transcription. The formation of such complexes releases a DNA strand which is difficult to purify away from RNA. Methods are described for preparing pure RNA strands.

Synthetic polymers are found to hybridize to strands that are not complementary in a Watson-Crick sense. Based upon mixing curve stoichiometry, CD, chemical and photochemical experiments one class of hybrids appear to be duplexes containing extrahelical bases. These structures can be illustrated schematically as:

```
- GGAGGA -
  ......
- CCTCCT -
   \/ \/
   T  T
```

A second class of hybrids appear to contain mismatches such as A·C, G·T and pyrimidine-pyrimidine oppositions. For example:

```
- GGAGGA -
  ......
- CTTCTT -
```

The physical properties of some of these hybrids have been examined by T_m, buoyant density, NMR, calorimetry and fluorescence measurements. Thermodynamic parameters for these and the corresponding perfect duplexes are presented. Hybrids containing extrahelical bases

have physical properties that may be relevant to frameshift mutagenesis.

Other studies concerned technical problems associated with replication of synthetic DNA's. The antibiotic netropsin can be added during replication of synthetic polymers in order to inhibit de novo synthesis of poly d(AT) but addition of netropsin then encourages de novo synthesis of poly d(TC)·poly d(GA). A method for the specific inhibition of de novo poly d(AT) synthesis using the synthetic antibiotic TANDEM is described. TANDEM also displays remarkable binding specificity apparently discriminating between base orientations in poly d(TAC)·poly d(GTA) and poly d(TCA)·poly d(TGA).

Acknowledgements

Above all I would like to thank Dr. A.R. Morgan for his help and advice and for providing an environment in which both curiosity and individual responsibility are encouraged. Such supervisors are rare indeed. Thanks also to Dr. Jeremy Lee for much sound advice and many happy hours of argument and to Ms. Denise Dombroski for patiently preparing vast amounts of DNA polymerase and other materials. Denise, Jeremy and Dick made our laboratory a pleasant place to work in for which I am grateful.

Many people generously made available their time and instrumentation. Mr. K. Oikawa recorded CD spectra on a spectropolarimeter belonging to Dr. C.M. Kay, Drs. M. Brauer and B.D. Sykes recorded NMR spectra, and Mr. V. Ledsham operated the analytical ultracentrifuges. Dr. R. Lewis and Ms. N. Mak provided advice on the operation of a scanning calorimeter belonging to Dr. R.N. McElhaney. Mr. D. Bacon provided instructions in the operation of the MMS computer graphics system in the laboratory of Dr. M.N.G. James. My thanks also to Drs. R.N. McElhaney, V. Paetkau, C. Bleackley, J. Silvius and M.O. Eze for encouraging an initial interest in research. The typing skills and endless patience of Ms. Georgette Jolicoeur is much appreciated.

Financial support in the form of Graduate Studentships from MRC Canada and an AHFMR research allowance is also gratefully acknowledged.

Table of Contents

Chapter		Page
I.	Introduction	1
	The Origins of this Project	1
	B-DNA and Its Isomorphs	3
	Nomenclature	3
	Fibre Diffraction	4
	The Crystal Structure of B-DNA	5
	Helical Polynucleotides Containing Non-Watson-Crick Base-Pairs	7
	Base-Pairing in Crystals	8
	Base-Pairing Specificity	8
	Non-Watson-Crick Helices	8
	Forces Favouring the Formation of Ordered Structures	16
	Hydrogen Bonding	20
	van der Waals-London Forces	20
	Base Stacking	21
	Other Aspects of Conformation	22
	The "Non-Rigid Nucleotide"	22
	Unusual Conformations in tRNA's	23
	Dynamic Fluctuations in Helices	28
	A-B Junctions	28
	Unusual Structures in Palindromic DNA's	29
	Conclusion	29
II.	Methods and Materials	32
	Chemicals	32
	Buffers, Salts and Reagents	32
	Commercial Polynculeotides	32
	DNA Binding Drugs	32
	Polymerase Substrates	33
	Radionucleotides and other Radioisotopes	33
	Chromatography	33
	Affinity Dyes	33
	DBAE-Cellulose	33
	Exonuclease I Chromatography	34
	Gel Exclusion Chromatography	34
	Pyrimidine Dimers	34
	RNA Polymerase Chromatography	34
	Thin-Layer Chromatography	34

Enzymes..... 35
DNA Polymerase..... 35
Endonucleases..... 35
Exonuclease I..... 36
Exonuclease III..... 36
Human Placental DNA Binding Protein..... 36
RNA Polymerase..... 36
Other Enzymes..... 36

Iodine Labelling..... 36

Physical Studies..... 37
Absorbance and T_m Measurements..... 37
Buoyant Density..... 37
Calorimetry..... 38
Circular Dichroism..... 38
Electrophoresis..... 39
Fluorescence Assays..... 39
Nuclear Magnetic Resonance..... 40
Nucleoside Analysis..... 40
Scintillation Counting..... 40
Sedimentation Velocity Measurements..... 41
Ultraviolet Dosimetry..... 41

Preparation of Double Stranded Synthetic DNA's..... 42
Reaction Conditions..... 42
Purification..... 42
Characterization..... 42

Preparation of Single Stranded Synthetic DNA's..... 44
Alkaline Buoyant Density..... 44
Polypyrimidine DNA's by Depurination..... 44
Polypurine DNA's..... 48

Preparation of Polypurine Synthetic RNA's..... 48

Preparation of [^{3}H] Labelled T7 DNA..... 50

III. Specific Inhibitors of DNA Polymerase I Catalysed De Novo DNA Synthesis..... 52

Introduction..... 52
De Novo Synthesis..... 52

De Novo Synthesis of Poly d(TC)·poly d(GA)..... 53

Netropsin Binding to Poly d(TTC)·poly d(GAA)..... 55

TANDEM - A Selective Inhibitor of Poly d(AT) Synthesis... 56
TANDEM..... 56
Effect of TANDEM on poly d(AT) Synthesis..... 58
TANDEM Specificity..... 58

TANDEM Inhibition of De Novo Synthesis in unprimed Reactions..... 63
Discussion - TANDEM Binding Specificity..... 67

Conclusion..... 68

IV. Polypurine RNA Self-Structure..... 69

Introduction..... 69

The Synthesis of Polypurine RNA's..... 69

Purification of Synthetic RNA's..... 74

Studies on Polypurine RNA's..... 77

Attempts to Form a Poly A Tetraplex..... 79

Discussion..... 82

Conclusion..... 84

V. Hybrid Polynucleotide Structures Containing Extrahelical Bases..... 85

Introduction..... 85

Hybrids Composed of Poly d(GGA) and poly d(TC)..... 86

Other Hybrids Containing Extrahelical Bases..... 95
Poly d(GA)·poly d(TTC)..... 98
Poly r(GAGA)·poly d(TCC)..... 101
Poly r(GAGA)·poly d(TTT)..... 102

Discussion..... 105

Conclusion..... 109

VI. Hybrid Polynucleotide Structures Containing Non-Watson-Crick Base-Pairs..... 110

Introduction..... 110

Hybrids Containing G·T Pairs..... 112

Hybrids Containing A·C Pairs..... 120

Hybrids Containing Pyrimidine·Pyrimidine Oppositions.... 121

Other Hybrids of Uncertain Structure..... 121

Strands Which do not Anneal..... 121

Discussion..... 121

Conclusion..... 126

VII. Physical Studies..... 127

Buoyant Density Studies..... 127

Nuclear Magnetic Resonance..... 129

Calorimetry..... 131
Calorimetric Profiles..... 132
Calculation..... 135
Results - Ordinary DNA Duplexes..... 136
Results - Hybrid Polynucleotides..... 141
Measurement of Charge Density..... 143
Frameshift Mutagenesis..... 145

Drug Binding..... 145

Terbium Fluorescence..... 146

Conclusion..... 150

VIII. Biological Studies..... 152

Human Placental DNA Binding Protein..... 153

Endonuclease Specificity..... 154

RNA Polymerase..... 159

Conclusion..... 159

IX. Conclusions and Further Experiments..... 162

DNA Polymerase I..... 162

Drug Binding Specity..... 162

Hybrids Containing Extrahelical Bases..... 163

Hybrids Containing Base Mismatches..... 164

Duplex Polynucleotide Structure..... 165

Biological Aspects..... 165

Appendix 1 Monovalent Cation Effects on Polynucleotide Stability..... 167

Appendix 2 Assumptions in Calorimetry..... 174

Biblography..... 178

Index to Figures

Figure 1-1 Bond angle nomenclature..... 3

Figure 1-2 The two common sugar puckers..... 4

Figure 1-3 Packing in d(ATAT) crystals..... 6

Figure 1-4 Adenine self-pairing..... 9

Figure 1-5 Guanine self-pairing..... 10

Figure 1-6 Pyrimidine self-pairing..... 11

Figure 1-7 Adenine-thymine self-pairing..... 12

Figure 1-8 Purine-purine, cytosine-guanine and uracil-guanine pairing..... 13

Figure 1-9 Other structures..... 14

Figure 1-10 Other structures..... 15

Figure 1-11 Phosphodiester bonds found within the yeast $tRNA^{phe}$ anticodon loop..... 25

Figure 1-12 Extrahelical base formation in yeast $tRNA^{phe}$..... 26

Figure 1-13 G.U pairing in yeast $tRNA^{phe}$..... 27

Figure 1-14 Formation of cruciform structures at palindromic DNA sequences..... 29

Figure 2-1 Separation of poly d(TCC) from poly d(GGA) by alkaline CsCl centrifugation..... 45

Figure 2-2 Preparation of poly d(GAA) by exonuclease III digestion..... 49

Figure 2-3 DNase I digestion of poly d(GAA)·poly d(TTC), poly r(GAA) mixture 51

Figure 3-1 De novo synthesis by E.coli DNA polymerase I..... 54

Figure 3-2 Netropsin binding to poly d(TTC)·poly d(GAA)..... 57

Figure 3-3 The structure of TANDEM..... 59

Figure 3-4 Effect of TANDEM on poly d(AT) synthesis..... 60

Figure 3-5 Effect of TANDEM on poly d(TAC)·poly d(GTA) melting..... 62

Figure 3-6 Selective inhibition of poly d(AT) synthesis..... 64

Figure 3-7 Effect of TANDEM on de novo poly d(AT) synthesis. 66

Figure 4-1 The two isomorphous base-tetrads..... 70

Figure 4-2 Kinetics of transcription..... 71

Figure 4-3 Effect of RNA polymerse to DNA strand ratio on net RNA synthesis..... 73

Figure 4-4 Purification of poly r(GA) by DBAE cellulose chromatography..... 76

Figure 4-5 Thermal denaturation profiles..... 78

Figure 4-6 Poly r(GAA) CD spectra..... 80

Figure 4-7 Poly d(GA) and poly d(GGA) will not form a hybrid of mixed composition..... 83

Figure 5-1 Determination of strand stoichiometry by mixing curve analysis..... 89

Figure 5-2 A CPK model of poly d(GA)·poly d(TTC) hybrid..... 82

Figure 5-3 Evidence for the existence of extrahelical bases as shown by the production of unexpected pyrimidine dimers..... 92

Figure 5-4 Spectroscopic evidence for the existence of extrahelical bases..... 93

Figure 5-5 Determination of hyperchromicity during complex formation..... 94

Figure 5-6 CD evidence for the formation of a hybrid between poly d(GA) and poly d(TTC)..... 99

Figure 5-7 Production of pyrimidine dimers by irradiation of poly d(TTC)..... 100

Figure 5-8 Reaction of poly r(GAGA)·poly d(TTT) with glyoxal. 103

Figure 5-9 Melting of poly d(GA)·poly d(TTC) hybrid in 0.40M NaCl, TE..... 107

Figure 6-1 The effect of NaCl concentration on poly d(GAA)· poly d(TTC) melting..... 116

Figure 6-2 Effect of strand ratio on relative hyperchromicities in 1.0 M NaCl, TE..... 117

Figure 6-3 UV mixing curve: poly dA and poly d(TTC)..... 120

Figure 6-4 Melting of the poly d(GGA)·poly d(CCC) hybrid in different concentrations of salt..... 121

Figure 7-1 Buoyant density shifts accompanying hybrid formation..... 128

Figure 7-2 NMR spectra of poly d(GGA)·poly d(TCC)..... 130

Figure 7-3 Melting of poly d(TC)·poly d(GA) followed by high sensitivity DSC..... 134

Figure 7-4 Errors associated with replicate calorimetric measurements..... 138

Figure 7-5 Enthalpies of helix formation in pyrimidine-purine DNA's..... 139

Figure 7-6 Effect of proflavine concentration on hybrid melting..... 147

Figure 7-7 Terbium fluorescence spectra..... 149

Figure 8-1 Human placental DNA binding protein assay..... 153

Figure 8-2 N.crassa endonuclease digestion of poly d(GA) hybrids..... 156

Figure 8-3 BAL 31 endonuclease digestion of poly d(GA) and poly d(GGA) hybrids..... 158

Figure 8-4 Synthesis of homopolymer RNA's off non-complementary single-stranded DNA's by RNA polymerase..... 160

Figure A-2(1) Effect of changes in partial specific heat capacities on apparent melting profiles..... 175

Figure A-2(2) Correction for residual single-strand stacking.... 176

Index to Tables

Table 1-1 Self-paired homopolynucleotide structures..... 17

Table 1-2 Non-Watson-Crick pairing between different homopolymers..... 18

Table 1-3 Heteropolymer complexes..... 19

Table 1-4 Stacking tendencies of sixteen ribonucleotides..... 24

Table 1-5 Bond rotations in yeast $tRNA^{phe}$..... 24

Table 2-1 Physical properties of synthetic pyrimidine-purine DNA's..... 43

Table 2-2 Properties of polypyrimidine DNA's..... 47

Table 3-1 Effect of TANDEM on the rate of synthesis of various DNA's..... 61

Table 3-2 The effect of TANDEM on the T_m of various synthetic DNA's..... 65

Table 4-1 Melting temperatures of polypurine nucleic acids..... 81

Table 5-1 RNA hybrids containing extrahelical bases..... 87

Table 5-2 Melting points of hybrids containing extrahelical bases..... 96

Table 5-3 Melting points of ordinary synthetic polynucleotides.. 97

Table 5-4 Probable structures of hybrid polymers containing extrahelical bases..... 104

Table 6-1 Non-Watson-Crick pairing in heteropolymer complexes... 111

Table 6-2 Probable structures of hybrid structures containing extrahelical bases..... 113

Table 6-3 Melting-points of hybrid polymer containing non-Watson-Crick base-pairs..... 114

Table 6-4 Hybrid polymers of uncertain structure..... 122

Table 6-5 Strands which failed to hybridize..... 122

Table 7-1 Thermodynamic properties of helix formation at $T = T_m$... 133

Table 7-2 Apparent thermodynamics of helix formation..... 141

Table 7-3 Apparent thermodynamics of defective helix formation... 142

Table 7-4 Calculated residue spacings..... 144

List of Abbreviations

A	adenine
Å	Ångstrom
b	axial charge spacing
BrA	8-bromoadenine
BrC	5-bromocytosine
BrU	5-bromouracil
C	cytosine
C^+	protonated cytosine
CD	circular dichroism
$\bar{C}_p$	partial specific heat capacity
CPK	Corey, Pauling, Koulton space-filling model
DBAE	dihydroxyborylaminoethyl cellulose
DNA	Deoxyribonucleic acid
dXMP	unspecified deoxyribonucleoside 5'-monophosphate dAMP deoxyadenosine monophosphate dCMP deoxycytidine monophosphate dGMP deoxyguanosine monophosphate dTMP deoxythymidine monophosphate
dXTP	unspecified deoxyribonucleoside 5'-triphosphate

EDTA	ethylenediaminetetraacetate
e	electron charge
G	guanine
ΔG	Gibbs free energy
G_{el}	electrostatic free energy
G_s	entropic contribution of counterion association
ΔH	enthalpy
I	inosine
J	joules
K	Boltzmann constant
KPi	potassium phosphate
m^6A	6-N-methyladenine
m_2^6A	6-N-dimethyladenine
mCi	millicuries
m^7G	7-methylguanine
M_r	relative molecular weight
n	cation valence
NMR	nuclear magnetic resonance
q	charge

r	charge separation
rh	relative humidity
R	gas constant
RNA	ribonucleic acid
rXTP	unspecified ribonucleoside 5'-triphosphate
S	entropy
SSC	saline sodium citrate
T	thymine
T	temperature
TANDEM	des-N-tetramethyl triostin A
TCA	trichloroacetic acid
T_m	melting temperature
Tris·HCl	hydrochloride salt of tris(hydroxymethyl)aminomethane
$\hat{TT}$	cyclobutane thymine dimer
U	uracil
$\hat{UT}$	cyclobutane uracil-thymine dimer
$\hat{UU}$	cyclobutane uracil dimer
W	work

UV	ultraviolet
ε	solvent dielectric constant
κ	Debye-Hückel screening parameter
ξ^{-1}	residual charge after counterion association
ψ	charge fraction neutralized by counterion association

Nomenclature

Normal Polynucleotides

i) Polynucleotides are indicated by the prefix "poly" [or alternatively by the subscript "n"]*.

ii) Ribo- and deoxyribopolymers are distinguished by the prefix "r" (for ribo) or "d" (deoxyribo). [It is also permissible to omit the specification in the case of ribopolymers but this can cause confusion.]

iii) The simple repeating sequence in synthetic polymers is indicated by surrounding brackets. These are omitted for homopolymers.

iv) The two complementary strands in a duplex are separated by a dot. The sequences in brackets are written assuming antiparallel strands.

Examples: A synthetic double-stranded DNA copolymer containing the simple sequence

```
- TTCTTCTTC -
  .........
- AAGAAGAAG -
```

is written as poly d(TTC)·poly d(GAA).

A synthetic double-stranded RNA copolymer containing the simple sequence

```
- UGUGUG -
  ......
- ACACAC -
```

is abbreviated poly r(UG)·poly r(CA). Note that the order in which strands are written is arbitrary and the sequence can be permuted.

* Square brackets indicate a convention found in the literature but but not used in this thesis.

For example poly r(AC)·poly r(GU) would also be correct.

The single-stranded DNA

- ACACAC -

is abbreviated poly d(AC) its single-strandedness being noted by the absence of a complementary strand.

Oligonucleotides

Oligonucleotides normally contain a defined sequence of known length therefore the whole sequence can be specified. Otherwise the nomenclature is the same as for polynucleotides.

Examples: The DNA oligonucleotide containing the following sequence

GTCAGTCA

would be abbreviated $d(GTCA)_2$.

A base-paired digonucleotide DNA duplex such as

AATTAATT
TTAATTAA

would be abbreviated $d(AATT)_2 \cdot d(AATT)_2$. The sequence is not permutable. Note this sequence causes some ambiguity because it could form an intramolecular hairpin structure:

AATT)
TTAA)

This would be referred to as a "$d(AATT)_2$ hairpin" which is distinctly different from $d(AATT)_2 \cdot d(AATT)_2$.

Unusual Base-Pairing Arrangements

The strand composition is indicated by specifying all strands. Since there may be more than one particular strand this is indicated by a numerical prefix.

Examples: Poly rG forms a four-stranded complex which would be abbreviated as 4 poly rG.

A triplex can be formed containing two strands of poly dT and one strand of poly dA. This can be abbreviated either as poly dT·poly dA·poly dT or 2 poly dT·poly dA.

Some polynucleotide structures contain protonated bases. This is indicated by a superscript "+" above the appropriate base. Poly rA forms such a duplex which would be written 2 poly rA^{+}.

Hybrid Nomenclature

In this thesis a number of hybrid polynucleotides are identified which contain strands in unusual stoichiometries. From these structures containing extrahelical base-pairs or non-Watson-Crick pairs are deduced. These hybrids are indicated by enclosing in brackets sufficient bases to account for the observed strand ratios.

Examples: Poly d(GGA) and poly d(TC) hybridize in a ratio that implies the hybrid contains four pyrimidine bases to three purines. This hybrid is then written poly d(GGA)·poly d(TCTC).

It is assumed that Watson-Crick pairing takes place where possible and that structures contain antiparallel strands. The hybrid containing poly d(GT) and poly d(TC) is then written as poly d(GT)·poly d(TC).

Random Copolymers

Polynucleotides containing a random base sequence are indicated by enclosing the base composition in brackets separated by a comma. The ratio of bases should normally be specified elsewhere in the text.

Example: A random ribopolymer containing 20% A and 80% G is abbreviated as poly r(G,A).

Other details of polynucleotide nomenclature can be found in the Journal of Molecular Biology 55: 299-310 (1971).

I. Introduction

Experiments over the past 20 years have shown that polynucleotides are capable of forming a variety of structures of which the B-helix is only one of many. These structures range in complexity from simple, stacked single-stranded helices to complex structures as in transfer RNA's. Helices can be left or right handed and one, two, three or four stranded. Even in duplexes the base-pairing need not be of the Watson-Crick type and helices containing protonated bases can be stable at neutral pH's. This review will attempt to document that polynucleotide conformation is tremendously flexible and that a variety of peculiar structures can be formed that seemed initially improbable.

The Origins of this Project

Ultimately base-sequence determines polynucleotide structure and the effects sequence exerts on structure have long been noted. This derives from studies on polynucleotides containing one or more bases arrayed in a simple repeating sequence. Historically the first of these polymers were *de novo* products of purified DNA polymerases and others were later prepared for the purpose of elucidating the genetic code (Wells *et al*, 1967). An advantage of these "synthetic" polynucleotides is that the physical properties of a base-sequence are amplified by repetition making it possible to measure the effect of sequence on structure. This project originated while investigating the properties of structures formed by one particular group of these polynucleotides.

As is detailed later polymers containing all purines in one strand and pyrimidines in the other can form three-stranded helices.

For example adding poly dT to poly dA·poly dT produces the triplex poly dT·poly dA·poly dT (Riley, Maling and Chamberlin, 1966) and under acidic conditions triplexes containing protonated cytosines can also form (Lipsett, 1964). This phenomenon is not restricted to homopolymers since poly d(TC)·poly d(GA)·poly r($U\overset{+}{C}$) triplexes can be prepared (Morgan and Wells, 1968). Considering this precedent it is reasonable to assume that poly d(TC)·poly d(GA)·poly d($T\overset{+}{C}$) might also form but dismutation of poly d(TC)·poly d(GA) under acid conditions yielded a structure first thought to be four-stranded (Johnson and Morgan, 1978). Later, however, Dr. Jeremy Lee demonstrated that dismutation of poly d(TC)·poly d(GA) does produce a triplex (Lee et al, 1979) and that the extra polypurine strand forms a second structure which is probably four-stranded. My first project was to see whether polypurine RNA's also form such structures and the results of these investigations make up chapter IV.

While preparing single-stranded polynucleotides for these studies it became apparent that single-strands that were not complementary in the Watson-Crick sense were still capable of hybridizing. Thus poly d(GA) not only hybridized to poly d(TC) but also hybridized to poly d(TTC) and poly d(TCC). The literature indicated that such phenomena had previously been observed in RNA's but not in DNA's or in polynucleotides of defined sequence. An investigation that started out of curiosity, rapidly expanded in scope and eventually formed the bulk of this thesis (chapters V-VIII).

Since considerable amounts of synthetic DNA were required for these studies a great deal of time was spent preparing various polynucleotides. This was done with E.coli DNA polymerase I which has

a disconcerting tendency to make poly d(AT) in preference to all other polymers. While attempting to specifically inhibit poly d(AT) synthesis a number of observations were made concerning polynucleotide structure, drug-binding specificity and polymerase function. Since these results may be of interest to anyone preparing synthetic polynucleotides they are presented as chapter III.

B-DNA and Its Isomorphs: Double stranded polynucleotides containing Watson-Crick Base-pairs.

Nomenclature

In order to discuss polynucleotide structure it is first necessary to briefly review bond and conformational nomenclature. Several different systems have been described but the one used here appears to be gaining the widest acceptance (Seeman et al, 1976; Drew et al, 1981). Bonds are designated in alphabetical order starting from the P-O5' bond and proceeding in a 5'→3' direction (Figure 1-1).

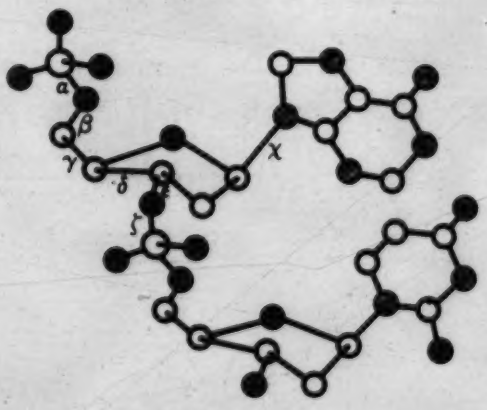


Figure 1-1 Bond angle nomenclature. The zero for bond rotation is defined by a cis-planar configuration and positive angles measured by clockwise rotation of the distal atom. The IUPAC-IUB zero for rotation about χ is defined by a plane formed by O-C1'-N9-C4 (purines) and O-C1'-N1-C2 (pyrimidines).

Sugar pucker is most properly described by a pseudorotation phase angle (Levitt and Warshel, 1978) and base rotation by a χ angle. Such sophistication is not normally necessary and terminology such as endo or exo and syn or anti usually suffices (Figure 1-2).

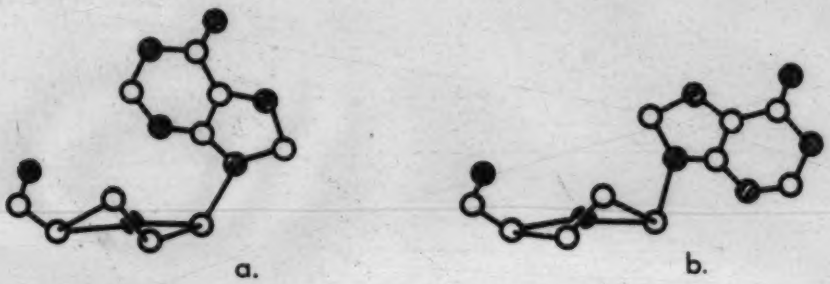


Figure 2-1 The two common sugar puckers and extreme χ rotations. (a) C3'-endo, syn. (b) C2'-endo, anti.

Fibre Diffraction

Early x-ray fibre diffraction studies suggested that three major forms of DNA helix exist depending on relative humidity (rh). These were designated A (75% rh., Fuller et al, 1965), B($\geq$ 95% rh., Langridge et al, 1960a,b) and C(66% rh., Marvin et al, 1961). Model building suggested that the B and C forms differ from A in sugar pucker. B and C were C2'-endo while A was C3'-endo. A appears to be the form preferred by double stranded RNA's, DNA·RNA hybrids (Milman, Chamberlin and Langridge, 1967) and dehydrated DNA's while B appears to be the solution form of DNA. C differs from B primarily in pitch but the biological relevance of this dehydrated Li salt remains uncertain. The 2'-hydroxyl seems to constrain RNA in an A-conformation regardless of humidity although this has recently been questioned (Zimmerman and Pheiffer, 1981).

Recognizing that such studies observe a sequentially averaged

structure synthetic DNA's with defined sequences have also been examined by fibre diffraction. The results have been relatively uninformative. Under conditions of high humidity B-DNA diffraction patterns are usually observed (Wells et al, 1977; Leslie et al, 1980) although the ability to interconvert between fibre forms does appear to be sequence dependent. Dehydration of poly d(AT) or poly d(GC) causes a B→D rather than B→A transition. These D forms are distinguished by a dinucleotide rather than mononucleotide repeat. Other fibre varients have also been described [B', C', C", E and S'] (Leslie et al, 1980) but the information available in fibre diffraction patterns is too limited to determine unambiguously their structures.

The Crystal Structure of B-DNA

Fibre diffraction data is now being supplanted by high resolution crystallographic studies of chemically synthesized oligonucleotides. The structure of a B-form double-stranded DNA dodecamer $d(CGCGAATTCGCG)_2$ was recently published (Drew and Dickerson, 1981; Drew et al, 1981; Dickerson and Drew, 1981; Wing et al, 1980). Bearing in mind the unusual nature of this sequence which is the Eco RI restriction site imbedded in alternating GC base-pairs these studies provide the first clues regarding how specific sequences affect the local geometry of a DNA helix.

The two most distinctive features of this structure are a gentle bend induced by lattice forces (19° over 11 base pairs) and propellor twisted bases of the type seen earlier in transfer RNA's (Sussman et al, 1978). Specific sequence dependent effects include variations in helical twist angle, base-tilt and ordering of water in the minor groove. Some conclusions could also be made concerning helix

dynamics. Temperature factors were larger in the backbone than in the core suggesting motion of the type observed by solid state NMR (Di Verdi and Opella, 1981). Backbone torsion angles varied (± 10-35 degrees) and sugar pucker was primarily C2'-endo although the minor variants C1'-exo and O1'-endo were also seen. Helical twist varied from base-pair to base-pair (9-11.4 per turn) and encompassed the average value of 10.4-10.6 determined for B-DNA in solution (Wang, 1979; Rhodes and Klug, 1980).

Three other deoxyoligonucleotide structures were reported before that of the B-DNA dodecamer. These bore no resemblance to B-DNA. In a crystal of d(ATAT) Viswamitra et al (1978) observed a structure in which a trans-gauche conformation about ζ and α generated an extended structure divided into two "mini-helices" base-paired to symmetry related molecules.

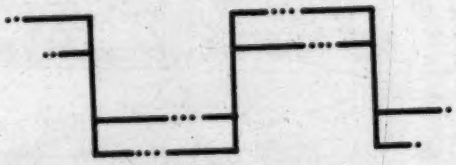


Figure 1-3 Packing in d(ATAT) crystals.

In another study d(CGCGCG) was observed to crystalize into a left handed helix (Wang et al, 1979). Pyrimidines were C2'-endo anti while purines were C3'-endo syn except for the 3' terminal G's which were C2'-endo syn. The alternating syn-anti conformation caused the backbone to zig-zag hence the name Z-DNA. A similar structure has been seen in d(CGCG) (Drew et al, 1980) and named Z'-DNA because of a slightly different sugar pucker.

These observations of Z-DNA explain the observation that when the salt concentration is raised the CD spectrum of poly d(GC) inverts almost completely (Pohl and Jovin, 1972). Z-DNA only forms in high salt (> 2M) because it is a slimmer helix containing a higher charge density.

Helical Polynucleotides Containing Non-Watson-Crick Base-Pairs

Experimental observations suggest that within the limits imposed by Watson-Crick base pairing a wide variety of double stranded helices can form. Other experimental data indicate that other base-pairing schemes are possible and that helices containing these odd pairs can also form.

The idea of specific base pairing between A and T(U) and G and C (Watson and Crick, 1953 ; Pauling and Corey, 1956) was formulated primarily on geometrical grounds. Hydrogen bonds stabilized these pairs and their isomorphism allowed them to form regular double-stranded helices with base-pairs in any orientation. This hypothesis prompted many attempts to demonstrate Watson-Crick base-pairing in crystals since the ~3.5 Å resolution (Donohue, 1969) of x-ray fibre diffraction was, and remains, insufficient for such purposes.

Base-Pairing in Crystals

Watson-Crick pairing in co-crystals of guanine and cytosine was first reported by O'Brian (1963) but only a variety of non-Watson base-pairs were observed in co-crystals of adenine and thymine or uracil. These structures included Hoogsteen pairs (Hoogsteen, 1959), reverse Watson-Crick (Sakore et al, 1969), reverse Hoogsteen (Katz et al, 1965) and a variety of self-paired arrangements. A·T(U) Watson-

Crick base-pairing has, in fact, never been seen in co-crystals of the bases on nucleosides. The first high resolution cyrstallographic demonstration was in a self-paired ApU·ApU structure (Seeman et al, 1976).

These and other types of base pairs are summarized in Figures 1-4 to 1-10. Examples are taken from a variety of published crystal structures including drug complexes and tRNA's (Voet and Rich, 1970).

Base-Pairing Specificity

A number of solution studies have been undertaken in an attempt to account for this multiplicity of base-pairing schemes. Normally pairing between bases and their derivatives can only be seen in apolar media since hydrogen-bonding solvents disrupt hydrogen bonding between bases. Typically $CDCl_3$ is used although G·C pairing is stable enough to occur in DMSO . Infrared spectroscopy (Kyogoku et al, 1966; Kyogoku et al, 1967; Nagel and Hanlon, 1972a,b) suggests that association constants fall in the order G·C ($10^4 - 10^5$) > G·G ($10^3 - 10^4$) >> A·U (10^2) > others (Kyogoku et al, 1969) the others only being detectable at relatively high base concentrations. NMR spectroscopy has shown this G·C pairing to be of the Watson-Crick type (Katz and Penman, 1966) but its not clear whether the A·U pair is of the Watson-Crick or Hoogsteen type. Indirect evidence suggests that these A·U solutions contain a mixture of both Watson-Crick and Hoogsteen pairs and that these two sites on adenine have approximately the same affinity for uracil (Nagel and Hanlon, 1972a,b). This readily leads to formation of A·U·A triples (Figure 1-9c) under some conditions.

Non-Watson-Crick Helices

Non-classical base-pairing is not limited to crystals and solutions of bases. Oligo and polynucleotides form complexes which,

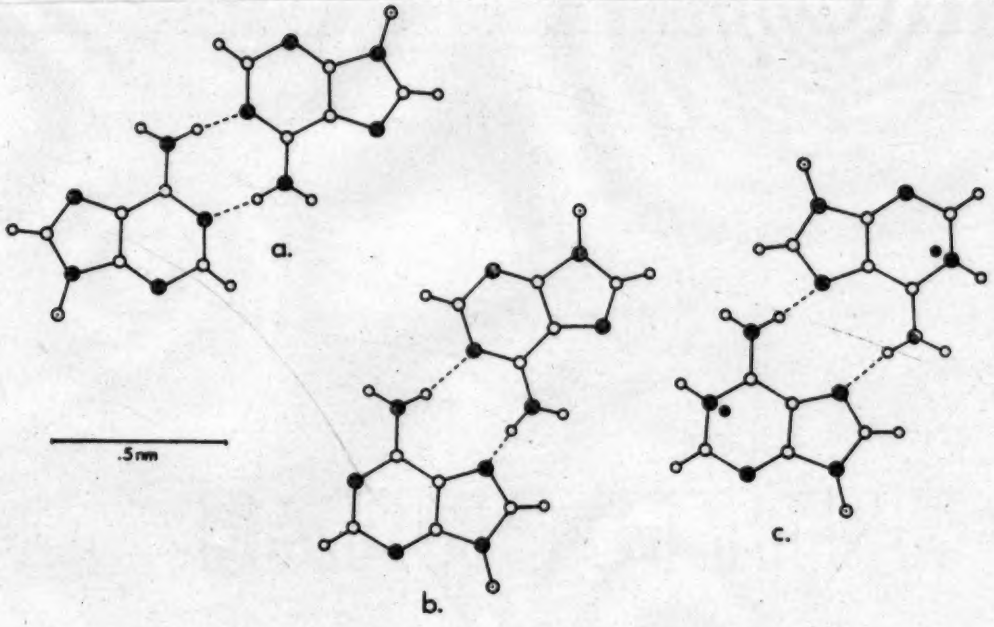


Figure 1-4 Adenine self pairing as observed crystallographically: (a) ApU proflavine complex (Westhof et al, 1980), (b) deoxyadenosine monohydrate (Watson et al, 1965), (c) adenine hydrochloride (Broomhead, 1948) and in tRNA (presumably neutral) (Quigley and Rich, 1976).

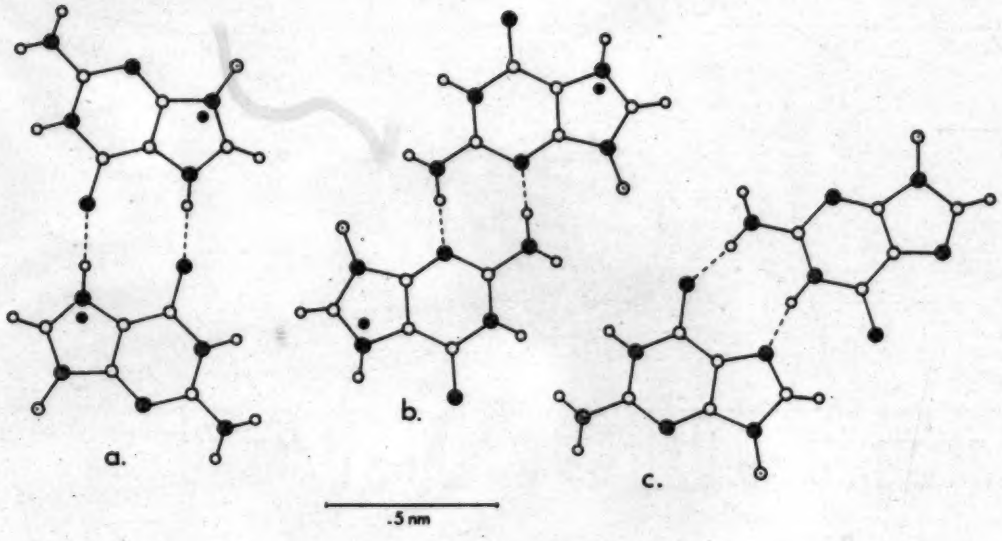


Figure 1-5 Guanine self-pairing as observed crystallographically: (a) guanine hydrochloride (Broomhead, 1951), (b) guanine hydrochloride-second pairing scheme (Broomhead, 1961), (c) guanine (Bugg et al, 1968) and tRNA (Quigley and Rich, 1976).

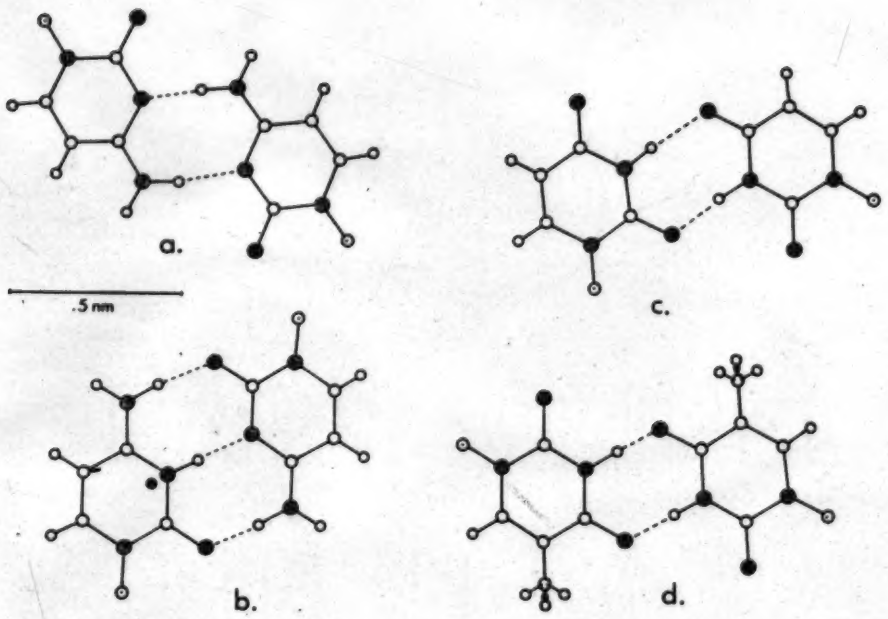


Figure 1-6 Pyrimidine self-pairing as observed crystallographically: (a) cytosine (Barker and Marsh, 1964), (b) $C^{+}\cdot C$ proflavine complex (Westhof et al, 1981), (c) UpA structure (Rubin et al, 1971), (d) thymine monohydrate (Gerdil, 1961) and uracil (Parry, 1954).

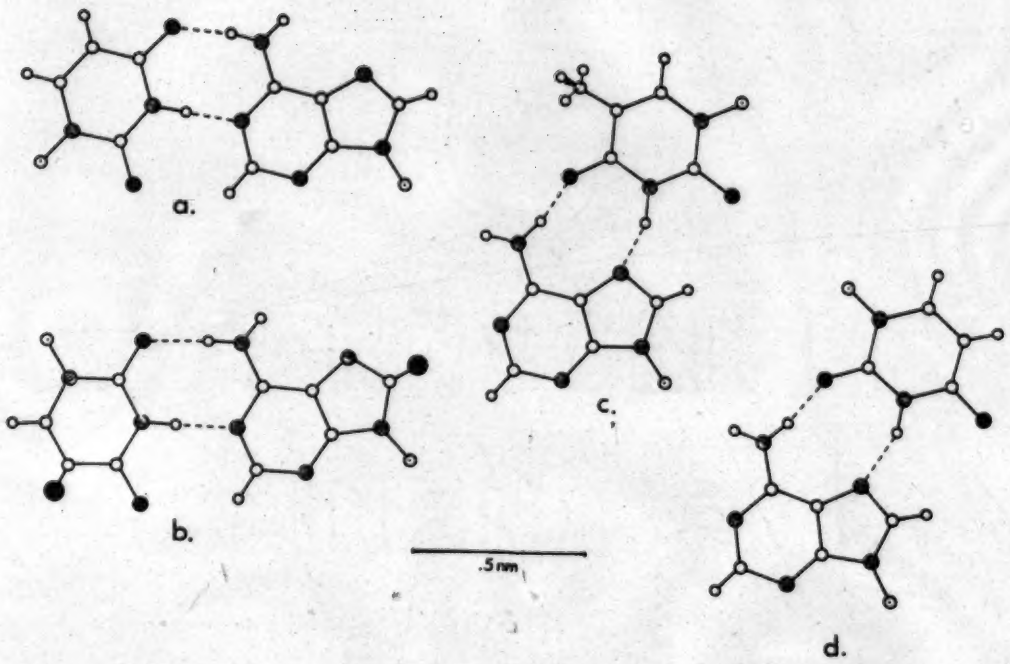


Figure 1-7 Adenine-thymine (or uracil) pairing observed crystallographically: (a) ApU structure (Seeman et al, 1976) and A·T pairing (Viswamitra et al, 1978), (b) BrU·BrA reverse Watson-Crick (Travale et al, 1969) and T·BrA (Travale et al, 1969), (c) T·A Hoogsteen pairing (Hoogsteen, 1969) and U·A (Mathews and Rich, 1964), (d) BrU·A reverse-Hoogsteen (Haschemeyer and Sobel, 1963) and U·A in tRNA (Quigley and Rich, 1976).

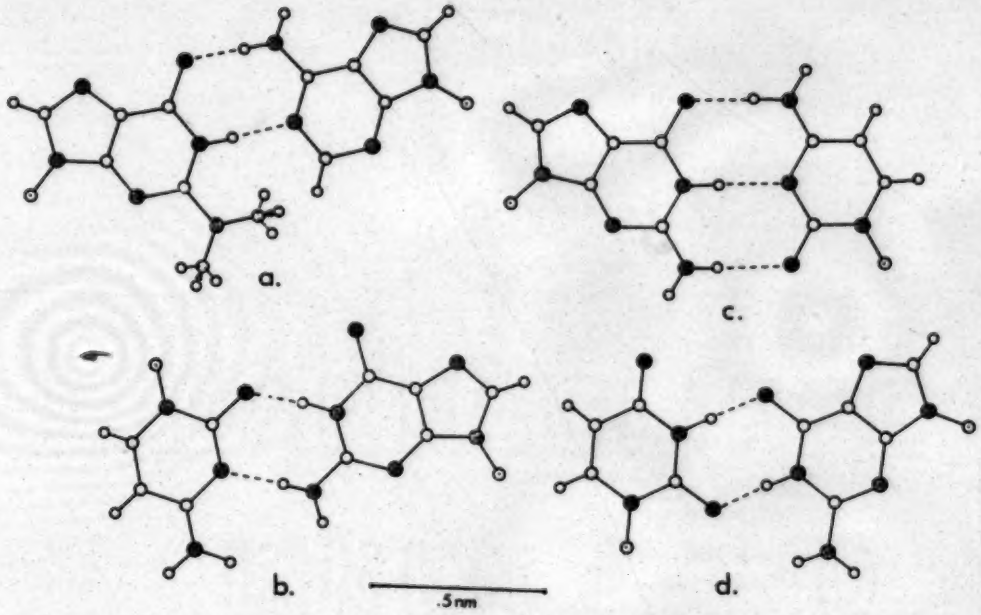


Figure 1-8 Purine-purine, cytosine-guanine and uracil-guanine pairing observed crystallographically: (a) A·m_2^2G in tRNA (Quigley and Rich, 1976) analogous to I·A "wobble" (Crick, 1966), (b) G·C reverse-Watson-Crick in tRNA (Quigley and Rich, 1976), (c) G·C Watson-Crick (O'Brian, 1963), (d) G·U "wobble" in tRNA (Quigley and Rich, 1976).

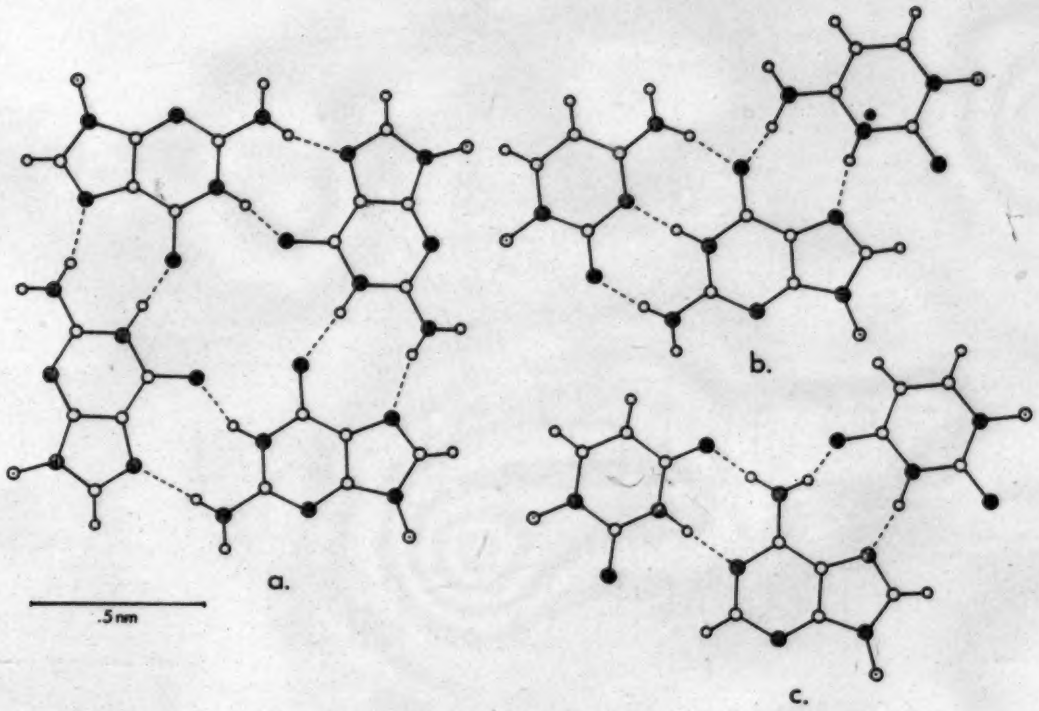


Figure 1-9 Structures inferred to exist from other than crystallographic data: (a)poly G and poly I (Zimmerman et al, 1975), (b) C·G·C^+ triplex (Lee et al, 1979), (c) U·A·U or T·A·T triplex (Riley et al, 1966).

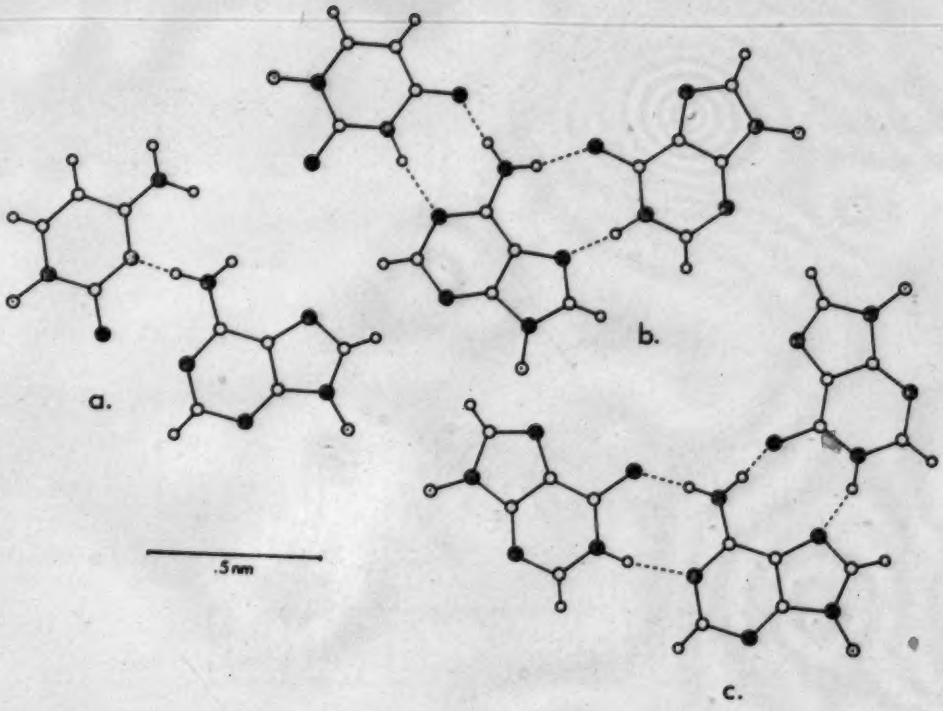


Figure 1-10 Structures inferred to exist from other than crystallographic data: (a) C·A "wobble" analogous to G·T wobble (Crick, 1966), (b) U·A·I triples (Declerq et al, 1975), (c) I·A·I triplex (Arnott and Bond, 1973).

although the detailed structure may be uncertain, cannot be due to Watson-Crick pairing. Homopolymer complexes show a variety of pairing schemes some of which are summarized in Tables 1-1 and 1-2.

Pairing between different homopolynucleotides often reflects the tendency to form Hoogsteen-pairs. This, along with Watson-Crick pairing, results in the formation of three stranded complexes (Table 1-2). Heteropolymers can also form non-Watson-Crick helices (Table 1-3) but the formation of regular helices requires that the base-pairs be isomorphous. This restricts the number of structures that can form.

Finally a few non-Watson-Crick base-pairs can coexist with Watson-Crick pairs in heteropolymer duplexes. In these experiments two polymer strands are hybridized one of which contains small numbers of non-complementary bases. Mixing curve stoichiometry makes it possible to determine whether the non-complementary bases are incorporated into the helix. The type of pairing occuring in these structures is uncertain. It may be of the "wobble" type (Crick, 1966) [Figures 1-8 (d), 1-10(a)] or may involve tautomeric bases. In some cases the non-complementary bases do not form base-pairs but are forced outside of the helix. These are often referred to as "extrahelical" bases (Fresco and Alberts, 1960) and can be indicated schematically as:

```
- UUUUU -
  .....
- AAAAA -
   \/
    U
```

Helices containing extrahelical and mismatched bases are discussed in greater detail in chapters V and VI.

Forces Favouring the Formation of Ordered Structures

Single-strands anneal forming ordered structures and two

Table 1-1

Self-Paired Homopolynucleotide Structures

Polymer	Probable pairing scheme	Comments	References
2 poly rA^{+}	Fig. 1-4(c)	-fibre, acid form	Rich et al (1961)
2 poly rA	Fig. 1-4(c)	-presumptive neutral form	Finch and Klug (1969)
2 poly $dC^{\frac{1}{2}+}$	Fig. 1-6(b)	-hemiprotonated acid acid form also poly poly $rC^{\frac{1}{2}+}$	Inman (1964)
2 poly rU	Fig. 1-6(c)	-NMR, very unstable	Young & Kallenback (1978)
4 poly rG	Fig. 1-9(a)	-symmetrical tetramer, also poly rI, extremely stable	Zimmerman et al, (1975)

Table 1-2

Non-Watson-Crick Pairing Between Different Homopolynucleotides

Complex	Probable pairing scheme	Comments	References
poly rA·2 poly rU	Fig. 1-10(c)	-also all DNA and mixed DNA·RNA hybrids	Riley et al (1966)
poly rC·poly rG ·poly rC^+	Fig. 1-10(b)	- as oligonucleotide narrow pH range (5-6)	Lipsett (1964)
poly rC·2 poly rG		-pairing uncertain as oligonucleotides	Lipsett (1964)
poly rA·2 poly rI	Fig. 1-9(c)	-other pairing possible	Howard & Miles (1977)
poly rU·poly rA ·poly rI	Fig. 1-9(b)		Declerq et al (1975)
poly rC·2 poly rI		-pairing uncertain	Chamberlin & Patterson (1965)

Table 1-3

Heteropolymer Complexes

Polymer	Probable pairing scheme	References
poly d(TC)·poly d(GA)·poly r(U$\overset{+}{C}$)*	Fig. 1-9(b) Fig. 1-9(c)	Morgan & Wells (1968)
poly d(TC)·poly d(GA)·poly d(T$\overset{+}{C}$)*	Fig. 1-9(b) Fig. 1-9(c)	Lee et al (1979)
4 poly d(GA), 4 poly r(GA)*	Fig. 4-1	Lee et al (1980)
2 poly d($\overset{+}{C}$T)	?	Gray et al (1980)
2 poly d(GT)	Fig. 1-8(d)	Early et al (1978)
2 poly r(GU)	Fig. 1-8(d)	Gray & Ratliff (1977)

* Note: In general repeating trinucleotide pyrimidine·purine DNA's and RNA's also form similar structures.

important driving forces are hydrogen-bonding and base-stacking interactions.

Hydrogen Bonding

Hydrogen bonding is not easily defined. The simplest description is that it involves an electrostatic interaction between a hydrogen attached to one electronegative atom and a second electronegative atom (Cantor and Schimmel, 1980) while more complete descriptions also include covalent interactions such as resonance (Jencks, 1969; Hamilton and Ibers, 1968). The strength of a hydrogen bond varies being dependent upon the geometry and types of atoms involved in the bond. Furthermore it cannot easily be measured since it is difficult to separate the "intrinsic" strength of a hydrogen-bond from other non-bonded interactions. A modern estimate suggest 0.0-0.4 Kcal/mol bonds (Levitt, 1978). Earlier estimates were higher (~ 1 Kcal/mol bonds) and on this basis the superior stability of G·C base-pairs was rationalized by the presence of one extra hydrogen bond (Crothers and Zimm, 1964).

Prior to 1964 doubts had already been expressed as to whether hydrogen-bonding played an important role in stabilizing polynucleotide helices (Ts'o, Melvin and Olson, 1963). Hydrogen bonding between bases can be completely disrupted by water for example and there is no correlation between hydrogen-bonding capacity and the denaturing ability of various compounds (Hanlon, 1966). These observations have led to a reevaluation of the hydrogen-bonding hypothesis.

van der Waals-London Forces

The specificity of Watson-Crick base-pairing cannot be explained in terms of hydrogen-bonding or rationalized on geometrical

grounds (Donohue and Trueblood, 1960). Voet and Rich (1970) have argued that the specificity and strength of pairing observed in solution at low base concentrations reflects "electronic complementarity" rather than simple geometrical complementarity. Thus G pairs with C in a Watson-Crick manner not simply because the hydrogen-bonding "fits" but rather because there is an underlying electronic component to this pairing. The hypothesis is also applicable to A·T(U) pairing but if A and U(T) are electronically "compatible" its not strictly in the Watson-Crick sense. Electronic compatibility is more properly a reflection of the fact that certain base geometries maximize van der Waals-London interactions and in non-polar, non-hydrogen-bonding solvents these interactions are further stabilized by hydrogen-bonding between the bases. This suggests that it is the so-called "weak" forces that are primarily responsible for the stability of Watson-Crick (and other) helices.

Base Stacking

The previous discussion perhaps over emphasizes the importance of base-pairing on structure. While base-pairing makes possible interactions between helices another factor stabilizing these structures is base-stacking. Stacking can be described as a strong tendency for bases to associate in vertical arrays and likely stabilizes double and multistranded structures by orienting bases so that interhelical pairing can take place. The effect is apparent in B-DNA structures for example. Base-pairs stack in van der Waals contact with one another excluding solvent from the helix core. The effect is readily seen in single-stranded polynucleotides which still form helical structures with base planes roughly perpendicular to the helix axis (Arnott et al, 1976).

The strength of base-stacking forces should be bourne in mind when attempting to predict polynucleotide structure. At present controversy still surrounds the relative importance of base-pairing and base-stacking in determining double or multistranded helix stability although recent calorimetric estimtes suggest that greater than half the enthalpy of helix formation might derive from stacking interactions (Filimonov and Privalov, 1978; Frier et al, 1981). NMR derived stacking tendencies agree with the limited number of values derived by other methods. Deoxyribonucleosides show similar stacking tendencies except that thymidine stacks more like cytidine than uracil (Solie and Schellman, 1968). Table 14 indicates the relative stacking tendencies of the sixteen ribodinucleotides.

Other Aspects of Conformation

The "Non-Rigid" Nucleotide

The ridged nucleotide hypothesis proposes that polynculeotides are composed of rigid nucleotide subunits and that conformational flexibility resides primarily in the internucleotide P-O and glycosidic (χ) bonds (Sundaralingam , 1975). Sugars are fixed in C3'-endo or C2'-endo conformations. Evidence for this hypothesis derives from x-ray crystallographic and NMR studies of mono and dinucleotides. These showed a strong preference for two conformations in ribonucletides C2'-endo, $\chi = 180°$, $\gamma = 60°$ and C3'-endo, $\chi = 180°$, $\gamma = 60°$.

The validity of the rigid nucleotide hypothesis is dubious. The terminology is incorrect since it is known that DNA is a dynamic structure with large, fast, coupled internal motions among the nucleotides (Hogan and Jardetsky, 1980) and within the phosphodiester backbone (Shindo et al, 1980; Di Verdi and Opella, 1981). There is also

little justification in stating categorically that one sugar conformation is to be preferred over another considering the small energy differences separating furanose conformational states (Levitt and Warshel, 1978). While the rigid nucleotides hypothesis is correct in that nucleotides prefer certain conformations there is little evidence that these preferences particularly restrict polynucleotide structure.

Unusual Conformations in tRNA's

Refined yeast $tRNA^{phe}$ structures have been reported to 2.5 Å and 2.7 Å resolution in which it is possible to observe many unusual aspects of polynucleotide conformation (Quigley and Rich, 1976; Jack et al, 1976; Holbrook et al, 1978). Although most of the tertiary interactions involve and may be dependent on the 2'-hydroxyl group, secondary interactions and backbone conformations are relevant to both DNA and RNA structure. It should be noted in the discussion that follows that at 2.5-2.7 Å resolution uncertainty still remains in fitting residues to electron density maps. In particular hydrogen bonding patterns and sugar conformation cannot be observed directly but must be inferred from backbone and base orientation.

Yeast $tRNA^{phe}$ is an L-shaped molecule composed of two irregular A-RNA helices at right angles to one another. Bond conformational distributions show some restriction of rotation about β, δ and ε but α, γ and ζ torsion angles are distributed about a full 360°. χ distribution is almost entirely anti ($\chi = 180°$) and sugar pucker is predominantly (~ 80%) C3'-endo (Holbrook et al, 1978). Many of the unusual bond rotations are associated with loops and bends in the molecule (Table 1-5). Within the anti-codon loop a sharp bend at U33 completely reverses the backbone direction (Figure 1-11). A complex

Table 1-4

Stacking Tendencies of Sixteen Ribodinucleotides at 20° in Water as Determined by NMR Spectroscopy

Strong	> Intermediate	> Medium	> Weak
GpG*	CpC	GpU	UpC
GpC	ApU	ApG	UpA
ApA	CpU	CpG	UpG
ApC	GpA	CpA	UpU

* Guanine stacking is difficult to measure due to formation of complex gels (Gellert et al, 1962) it is normally assumed to be extremely strong. Taken from (Lee et al, 1976; Ezra et al, 1977) arbitrarily divided into four categories.

Table 1-5

Bond Rotation (in degrees) observed at unusual $tRNA^{phe}$ Chain Configurations*.

Bond	Reverse Bend		Extrahelical U			ARNA (fibre)‡	tRNA (median)
	A35	G3	M^7G46	U47	C48		
α	200	240	230	90	250	280	280
β	250	90	70	190	220	210	160
γ	120	170	190	40	180	40	60
δ	80	80	180	140	140	80	80
ε	230	220	260	220	220	180	220
ζ	330	260	270	310	60	210	280
χ	190	180	260	260	230	---	170
pucker	C3'-endo	C3'-endo	C2'-endo	C2'-endo	C2'-endo	C3'-endo	C3'-endo

* From data tabulated in Sussman et al (1978). Bond nomenclature and χ values have been altered to maintain consistency with the IUPAC-IUB nomenclature indicated earlier.

‡ From Arnott and Hukins (1972) nomenclature has been modified.

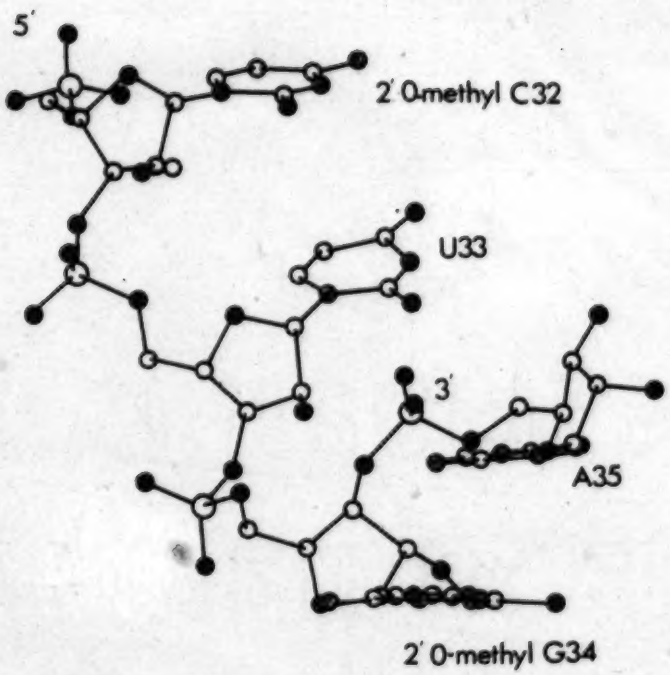


Figure 1-11 Phosphodiester bends found within the yeast $tRNA^{phe}$ anticodon loop. View is looking towards two of the anticodon bases.

pattern of hydrogen-bonding and van der Waals contacts between U33 and its neighbours appears to stabilize the structure (Quigley and Rich, 1976) and unusual bond rotations at residues A35 and A36 effect the turn (Table 1-4). Another unusual examples of backbone flexibility is found at U47. Here weak stacking between C48 and M^7G46 results in U47 being excluded into an extrahelical location (Figure 1-12) and again unusual bond rotations are responsible for this structure (Table 1-5). A notable perculiarity about this extrahelical residue are three consecutive C2'-endo sugars (Table 1-5). tRNA's also contain non-Watson-Crick base-pairs (Figures 1-4 to 1-10) an example of which is indicated in more detail in Figure 1-13. This is a G·U "wobble" pair predicted by Crick (1966) in order to rationalize flexibility in codon-anticodon recognition.

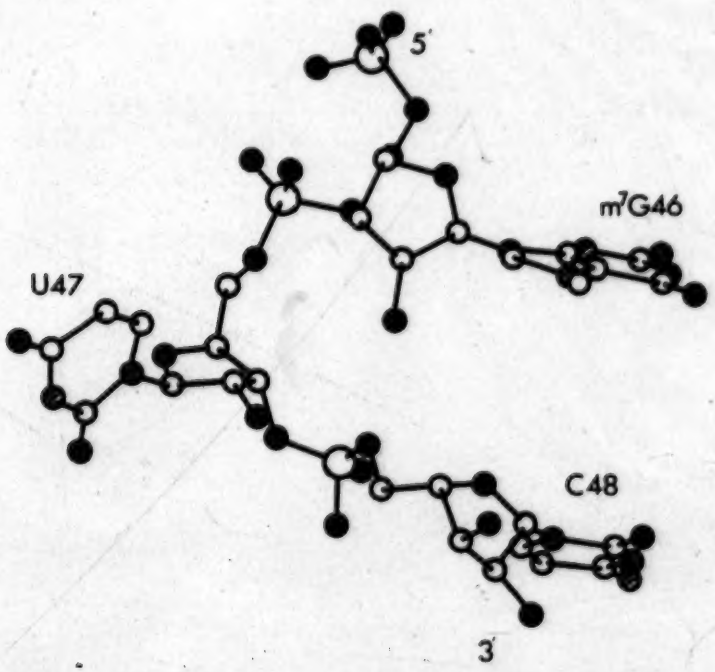


Figure 1-12 Extrahelical base formation in yeast $tRNA^{phe}$.

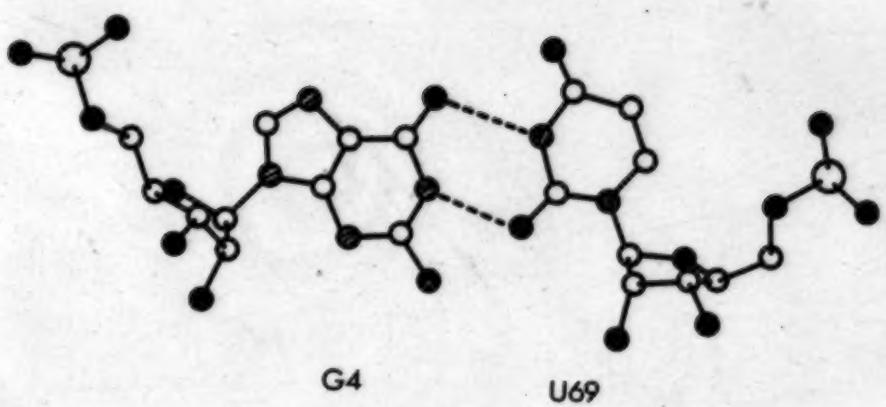


Figure 1-13 G·U pairings within the yeast $tRNA^{phe}$ amino acid acceptor arm.

Dynamic Fluctuations in Helices

Tritium exchange and NMR studies on natural DNA's in solution provide information concerning the dynamics of helices. Early et al (1981a,b) observed the relaxation behaviour of imino protons in small restriction fragments. Assuming that at elevated temperatures solvent exchange dominates relaxation measurments in short fragments the rate and van't Hoff activation energy of A·T base pair opening could be calculated. A rather high energy (~ 16 Kcal/mol) was deduced in agreement with stopped flow hydrogen exchange techniques (Mandal et al, 1979). A·T rates of proton exchange were approximately three times as high as the G·C rate. In another NMR study on somewhat larger non-homogeneous fragments Hogan and Jardetzky (1980) measured the amplitudes of motions with respect to the helix (z) axis. Within the geometrical limits of their measurements amplitudes of ±20-30° with time constants of $\sim 10^{-9}$ s^{-1} were observed. They concluded that such motions were at least consistent with rapid isomerization between a classical B conformation and a type of structure proposed by Levitt (1978) containing O1'-endo sugars, propeller twisted bases and altered twist angle. Such a structure is also very similar to the B-DNA dodecamer discussed earlier but one should note that this structure was refined using a Levitt energy minimization function so the agreement is not as coincidental as might first appear.

A-B Junctions

Selsing et al (1978) prepared a hybrid of poly dG and $[r(C)_{11}$ $d(C)_{16}]$ and examined the proton NMR spectra. Spectra were consistent with A and B conformations coexisting in the same duplex although a fortuitous equilibrium mixture of molecules could not be discounted.

Model building suggested a bend was required at the A-B junction in order to link the different helical types. These observations provide some support for the alternating B-DNA model of Klug et al (1979) which suggests that the unusual physical properties of poly d(AT) are due to an alternating C3'-endo(A), C2'-endo(T) structure such as would occur at an A-B junction.

Unusual Structures in Palindromic DNA's

Supercoiled double-stranded DNA's contain sites which are substrates for single-strand specific endonucleases (eg. Legerski et al, 1977). It has been shown that these sites are often palindromic DNA sequences and that these nucleases are nicking at potential hairpin terminii (Panayototas and Wells, 1981).

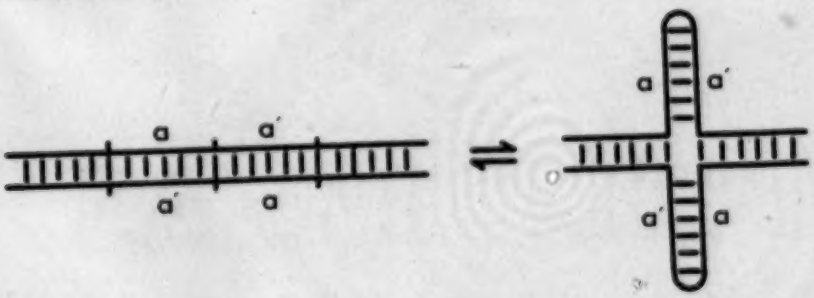


Figure 1-14 Formation of cruciform structures at palindromic DNA sequences. Experimentally a minimum of two bases appears necessary in order for unimolecular loop formation to be kinetically preferred over bimolecular duplex formation (Haasnoot et al, 1980). Model building suggests that surprisingly little distortion is required at the cross-over (Sigal and Alberts, 1972).

Conclusion

Considerable precedent exists for arguing that polynucleotide formation is highly flexible and that within stereochemical limits many structures can form. What is more difficult to evaluate is the

actual destabilizing energies associated with unusual polynucleotide structures. Polynucleotide stability represents a balance between favourable enthalpies of base-pairing and base stacking, unfavourable energies associated with counterion binding and charge repulsion (Record et al, 1978; 1981) and unfavourable energies associated with stereochemical distortion of the phosphodiester and (deoxy)ribose backbone.

The destabilizing effects of structures such as hairpin loops and extrahelical bases have been estimated from van't Hoff analysis of oligoribonucleotide melting. In moderate to high ionic strengths unfavourable free energies of $\Delta G^\circ = 3 \pm 1$ Kcal/mol extrahelical base and $\Delta G^\circ \geq 8 \pm 2$ Kcal/loop (single base loop containing 3 bases) are commonly quoted (Salser, 1977; Tinoco et al, 1973; Fink and Crothers, 1972). The loop values have been criticized as overestimates based on calorimetric measurements on tRNA melting (Privalov and Filimonov, 1978) but independently derived values for extrahelical base free energies are not available. Standard Watson-Crick base-pairs free energies of formation have been estimated to be 1.2-5 Kcal/mol base-pair, (depending on nearest neighbour and base-pair type) from oligoribonucleotide melting analysis (Tinoco et al, 1973) and 0.8-3.6 Kcal/mol base-pair on the basis of tRNA and poly rA·poly rU melting studies (Privalov and Filimonov, 1978). Consider, then, whether the following structure might form within a double-stranded polynucleotide helix.

Taking the maximum destabilizing free energy of an extrahelical base

as +3 Kcal/mol, the free energy of base-pairs formation as -5 Kcal/mol base-pair for a G·C pair adding onto a G·C pair and -2.2 Kcal/mol for the other two pairs this gives a free energy of formation at 25° of:

$$\Delta G^\circ = 3 - 2.2 - 2.2 - 5$$
$$= -6.4 \text{ Kcal/mol or}$$
$$= -0.8 \text{ Kcal/mol phosphate.}$$

Taking 1 Kcal as a minimum estimate for the stability of each base-pair this helix is just melting at 25°:

$$\Delta G^\circ = 3 - 1 - 1 - 1$$
$$= 0$$

Thus crude thermodynamic arguments predict that Watson-Crick helices can be "hosts" to some unusual structures. While it is not clear exactly how stable these structures might be their existence is to be expected.

This thesis concerns itself with investigating the physical properties of structures such as these containing extrahelical or mismatched base-pairs. A long term goal of this research is to investigate whether such structures are relevant to an understanding of the physical phenomena underlying such processes as DNA repair and recombination.

II. Methods and Materials

Chemicals

Buffers, Salts and Reagents

CsCl was from Pierce (Sequenol Grade) and $CsSO_4$ from Kerr-McGhee. Buffers were prepared and titrated as 1.0 M solutions and diluted as required. Polynucleotides were normally stored in 10 mM Tris·HCl pH8, 0.1 mM EDTA (TE buffer). Concentrated stock solutions of buffers and salts were normally millipore filtered (Millipore HA 0.45 μm) to remove dust and other insoluble residues. Solvents (ethanol, toluene, chloroform) were routinely redistilled, water was deionized and then distilled typical conductivities were <1 μmho. Other reagents were of reagent grade or better.

Glyoxal was prepared by refluxing solid Fisher glyoxal (about 3 hours) in distilled H_2O, neutralizing with concentrated NaOH (pH 7.0) and filtering off Norit to remove coloured impurities. Final concentrations were determined by density.

Commercial Polynucleotides

Poly dT, poly dC, poly dA, poly rA and poly r(A,G) were purhased from PL Biochemicals and stock solutions prepared by resuspending in TE buffer and heating 2' at 100°. Poly r(A,G) contained adenine and guanine in a 0.61:0.39 ratio. Final concentrations were determined spectrophotometrically.

DNA Binding Drugs

TANDEM (N-desmethyl-triostin A, $\varepsilon_{325} = 1.11 \times 10^4$) was a gift of Dr. R.K. Olsen (Utah State University), spermine-bis-acridine and netropsin ($\varepsilon_{296} = 2.02 \times 10^4$) gifts of Dr. J.H. Van de Sande (University of Calgary).

Polymerase Substrates

Ribo and deoxyribonucleoside triphosphates were purchased from PL Biochemicals and stock solutions stored frozen in TE. Base and nucleoside standards were purchased from a variety of sources (Sigma, PL Biochemicals, Raylo) and used without further purification. dBrCTP (a gift of Dr. D. Pulleyblank) was prepared by bromination of dCTP and dBrUTP (a gift of Ms. T. Elgert) by nitrous acid deamination of dBrCTP. Diaminopurine deoxynucleoside trisphosphate was a gift of Dr. J.H. Van de Sande.

Radionucleotides and Other Radioisotopes

Radioactive nucleoside triphosphates were purchased (NEN, Schwartz-Mann, Amersham, ICN) and used without further purification. A check on purity showed that some preparations did contain significant levels of di- and mono-phosphates. $Na^{125}I$ was purchased from the Edmonton Radio Pharmaceutical Centre and all other radioisotopically labelled materials from NEN.

Chromatography

Affinity Dyes. Malachite green affinity absorbant gels were purchased from Boehringer-Mannheim. Samples of DNA were loaded in 20 mM KPi, pH 6.9 and eluted with linear 0-2.0 M gradients of $NaClO_4$ in 20 mM KPi, pH 5.9 as described by the manufacturer. Columns were kept protected from light with aluminum foil and washed prior to use with 5.0 M $NaClO_4$, 20 mM KPi pH 5.9.

DBAE - Cellulose. Acetylated N - [N' - (m - dihydroxyborylphenyl) succinamyl] aminoethyl cellulose (DBAE-cellulose) was purchased from Collaborative Research. Samples were loaded in Buffer 1 (0.20 M NaCl, 0.01 M $MgCl_2$, 0.05 M Tris·HCl pH 8.7, 20% ethanol V/V) and washed, then

bound RNA eluted with Buffer 2 (0.20 M NaCl, 0.05 M NaOAc pH 5.0) as described by McCutchan et al (1975). All procedures were carried out at 4°.

Exonuclease I Chromatography. Exonuclease I was first passaged on a (0.75 x 25 cm) Sephadex G·100 (fine) column in 50 mM Tris·HCl pH 8.0, 10 mM EDTA, 5% glycerol, 0.1 M KCl and (when this didn't remove RNases) then on a (0.8 x 10 cm) Biogel A 0.5 m (200-400 mesh) column in the same buffer.

Gel Exclusion Chromatography. Double-stranded polynucleotides and RNA's were normally purified on small (0.5 x 25 cm) columns of A 1.5 m or A 5 m 50-100 mesh Biogel in TE. Columns were washed prior to use with a solution of 4 M KCl, 0.1 M NaOH. Single stranded polydeoxyribonucleotides were purified on the same columns in 5 mM NaOH, 0.1 mM EDTA to minimize non-specific losses.

Pyrimidine Dimers. Pyrimidine dimers were separated by descending paper chromatography on Whatman number I paper developed with n-butanol:acetic acid:H_2O (80:12:30) as described by Setlow and Carrier (1966). This solution had to be prepared fresh every 2-3 weeks due to ester formation.

RNA Polymerase Chromatography. Columns (DE-52, Biogel A 1.5 m (200-400 mesh), DEAE-sephadex) were prepared as described by the manufacturers and used as per published procedures (Burgess, 1969). DNA-cellulose was prepared according to Alberts and Herrick (1971) and used as described by Burgess and Jendrisak (1975).

Thin-Layer Chromatography. Purity of radioactive nucleotides was checked on Baker-Flex cellulose PEI plates in 20 M LiCl: 2.0 M HCOOH (1:1 V/V) or alternatively on Eastman 13255-cellulose plates developed in isobutyric acid: H_2O:NH_4OH (60:40:1) (Randerath and Randerath, 1967).

Production of nucleosides for nucleoside analysis was monitored on Baker-Flex cellulose PEI plates developed in TE.

Enzymes

DNA Polymerase. A variety of different polymerase preparations were used, prepared by slightly different methods and by a number of individuals (Ms. J. Forsythe, Ms. T. Elgert, Ms. D. Dombroski and Mr. R. Doerkson). Normally DNA polymerase I was prepared by the method of Jovin et al (1969) except that the order of chromatography was over DEAE cellulose, Sephadex G-100 and then phosphocellulose P-11 in order to recover exonuclease III as a G-100 by-product. DNA polymerase prepared using this method has specific activities of 17 to 900 U/mg (Richardson et al, 1964) and endonuclease activities of 0.3 to 200 U/mg*. More recently the order of purification has involved Sephadex G-50, DEAE-cellulose, phosphocellulose P-11 and an additional chromatography step on LKB HA-Ultragel (after Richardson et al, 1964). These preparations showed specific activities of circa 700 U/mg and essentially undetectable levels of endonuclease.

Endonucleases. DNase I (2000 U/mg, RNase free (Kunitz, 1950)) was purchased from Boehringer-Mannheim and N.crassa endonuclease from Miles (900 U/mg (Rabin and Frazer, 1970)). S_1 nuclease (0.5 U/mg (Vogt, 1973)) was a gift of Dr. D. Johnson. Pseudomonas BAL 31 endonuclease (Alteromonas espejiana endonuclease) a gift of Ms. J. Forsythe was prepared according to Gray et al (1975) and had a specific activity of 4800 U/mg (PM2 nicking in 0.5 M NaCl, 0.1 M Tris HCl pH 8, 1 mM Mg SO_4, 1 mM $CaCl_2$, 0.1 mg/ml gelatin).

* Defining 1 unit as the amount of enzyme required to put 1 nick per molecule (on average) in 1 OD unit of covalently closed circular PM2 DNA in 10min at 37° in 10 mM $MgCl_2$, 50 mM KPi pH 7.5, 0.1 mg/ml gelatin.

Exonuclease I. Exonuclease I (a gift of Ms. T. Elgert) was further purified by chromatography on Sephadex G-100 and Biogel 200-400 mesh A 0.5 m. Final specific activity was 2000 U/mg (Lehman and Nussbaum, 1964).

Exonuclease III. Exonuclease III (a gift of Ms. D. Dombroski) was recovered as a byproduct of DNA polymerase I preparation and chromatographed on phosphocellulose P-11. Final specific activity was circa 17000 U/mg (Richardson, Lehman and Kornberg, 1964).

Human Placental DNA Binding Protein. Human placental DNA binding protein was a generous gift of Dr. Ross Feldberg (Tufts University). Two preparations were used "Seph IV" and "HA-ultrogel" fractions. Binding protein was assayed by the ability to retain ultraviolet treated T7 DNA on nitrocellulose filters (Feldberg and Grossman, 1976).

RNA Polymerase. RNA polymerase was purified using a procedure of Mr. P. d'Obrenon based closely upon that of Burgess (1969). The product was found to be significantly contaminated with polynucleotide phosphorylase (Kimhi and Littauer, 1968) and for this reason an additional DNA-cellulose step (Burgess and Jendrisak, 1975) was employed. SDS gels showed all four subunit types present although no attempt was made to quantitate σ factor. The final product had a specific activity of 6.8 U/mg (Burgess, 1969). Traces of residual polynucleotide phosphorylase activity (Kimhi and Littauer, 1968) were inhibited by adding 0.2 mM KPi to all reactions.

Other Enzymes. Polynucleotide Kinase (T4) was a gift of Dr. J.H. Van de Sande, other enzymes were purchased commercially.

Iodine Labelling

Samples were labelled using a commercial solid-phase bound

lactoperoxidase-glucose oxidase system (Biorad "enzymobeads"). Proteins were desalted into 0.20 M KPi pH 7.5, 0.1 mM EDTA on 50-100 mesh Biogel P-100 columns and then incubated with NaI (1 μCi), glucose (0.4%) and Biobeads according to manufacturers specifications. Reactions were stopped with sodium metabisulfite (1 mM) and product purified away from reactants on precalibrated Biogel P-100 columns into 30 mM Tris pH 8.0, 0.2 mM EDTA, 13 mM β mercaptoethanol, 5% glycerol, 0.75 M NaCl. Labelling was approximately 3500 cpm ^{125}I per μg protein assuming 80% recovery of protein off each column.

Physical Studies

Absorbance and T_m Measurements. Absorbances were recorded on a Gilford Model 2400 spectrophotometer equipped with thermoplates and a Haake heating unit. Wavelength accuracy was checked with holmium oxide filters (± 0.5 nm) and absorbance accuracy against standard neutral density filters (± 2%). Thermal denaturation experiments were carried out in quartz microcells. Samples (250-300 μl) were over-layed with paraffin oil (Fischer) and tightly stoppered with Teflon caps. Weighing of cuvettes before and after melting showed negligible loss of solvent. Samples were degassed only if being heated to greater than 80°, corrections for solvent and cuvette expansion were not applied. The T_m was taken to be the midpoint of each thermal transition.

Absorbtion and difference spectra were recorded on a Beckman DU-8 spectrophotometer equipped with λ scanning software in the laboratory of Dr. R.N. McElhaney.

Buoyant Density. A Beckman model E analytical ultracentrifuage equipped with cell scanner-multiplexer accessory was used for all buoyant density measurement. Samples were centrifuged in 4° single-sector cells

in either 4 cell (AnF Ti) or 2 cell (AnH Ti) rotors at 48000 rpm at 20°. Calculations assumed $\omega^2/2s = 0.01792$ (Ifft, 1961) for CsCl solutions at this speed. Densities were calculated relative to an internal λ DNA standard ($\rho = 1.7028$ in CsCl) and checked when possible using the iso concentration method (Szybalski, 1968).

Preparative buoyant density experiments used a Beckman Ti 50 angle rotor. Samples (10.7 ml) were layered over with paraffin oil and centrifuged at 35000rpm, 66-90 hours in a Beckman L2-65B ultracentrifuge, punctured and fractions collected manually (Morgan et al, 1974). Densities were determined by refractometry against standard calibration curves.

Calorimetry. Melting enthalpies were recorded with a Microcal -1 adiabatic high sensitivity scanning microcalorimeter in the laboratory of Dr. R.N. McElhaney. Typical instrument settings were full scale range 30 μV, 2x x-axis sensitivity, 8x y-axis sensitivity, 1.20 mcal/min calibration pulse, 0.9 deg/min heating rate. Sample volume was 0.876 ml. Experiments typically required about 0.8 μmol DNA for good sensitivity. Melting enthalpies were determined by planimetry; errors in area analysis being ±3%. Normally melting enthalpies were recorded twice with samples being allowed to reanneal overnight by slowly cooling (~5%/hr) the calorimeter cells. Calculations are detailed later in chapter VII.

Circular Dichroism. CD spectra were recorded on a Cary 60 spectropolarimeter equipped with model 6001 CD accessory and thermostated cells in the laboratory of Dr. C.M. Kay. Spectral reduction was as described in Wells et al (1970). Non-interacting and theoretical spectra were measured by summing individual spectra manually and normalizing to a per residue basis. Details of these calculations are described in

other chapters. Samples were centrifuged prior to analysis in a bench-top clinical centrifuge to remove dust.

Electrophoresis. Agarose gel electrophoresis was normally performed in 0.6-1.6% agarose (Sigma type II) gels using the procedure of Mickel et al (1977) in a flat bed gel apparatus obtained from the Aguebogue machine and repair shop (Aguebogue, N.Y.). Molecular weight standards were Eco RI cut λ DNA (Thomas and Davis, 1975) and Hae III cut PRZ 2 DNA (Wells et al, 1979). Ethidium bromide stained gels were visualized with ultraviolet light and photographed with polaroid 667 film and a Polaroid CU-5 Land Camera equipped with orange (02) filter.

For nearest neighbour analyses samples were electrophoresed on Whatman Number 1 paper in a Savant flat-bed apparatus at 20 V/cm for 6 hours in 0.05 M sodium citrate pH 3.7 as described previously (Morgan, 1970).

Fluorescence Assays. Ethidium bromide fluorescence assays were performed on a Turner-430 spectrofluorimeter with excitation wavelength of 525 nm and emission at 600 nm and fully open (60 nm) slit widths as described earlier (Morgan et al, 1979a, b). The pH 8 buffer used here contains 0.5 μg/ml ethidium bromide, 5 mM Tris HCl pH 8.0, 0.5 mM EDTA and the pH 5 buffer 1 μg/ml ethidium bromide, 20 mM Na acetate pH 4.8, 0.2 M NaCl and 0.5 mM EDTA. On the x 100 scale of this instrument 10 μl of 1.0 A_{260} calf thymus DNA gives 70 fluorescence units at pH 8 and 30 at pH 5. Samples were prepared in 10 x 75 mm test tubes pretreated with 2% dichlorodimethyl silane (Sigma) in $CHCl_3$ to prevent non-specific adsorption. Two minutes in a 100° Temp-Block was used for heat denaturing samples.

Terbium fluorescence assays were carried out using the same

spectrofluorimeter. $TbCl_3 \cdot 6H_2O$ (Alfa Products) was prepared as a 0.1 M stock solution in 0.1 M HCl and $TbCl_3$ assay solutions prepared daily as described elsewhere. Fluorimeter settings varied but slit widths were normally set at minimum (15 nm) band widths in order to minimize spectral overlaps. Typical excitation and emission settings were 265 and 545 nm respectively. Quartz 1 cm fluorimeter cuvettes were used and samples prepared only in new, disposible 10 x 75 mm glass test tubes because old, washed test tubes used in ethidium fluorescence assays contained some unidentified material that severely interfered with Tb fluorescence measurements.

Nuclear Magnetic Resonance. Phosphorous 31 NMR spectra were kindly recorded by Drs. B.D. Sykes and M. Brauer on a 270 MHz (proton) Brüker NMR spectrometer. Fourier transform spectra were collected at 69.3 MHz at 300° K with instrument settings of sweep width ±2500 Hz, 4096 points, 23 µsec pulse width (~76°) and 1 sec delay between acquisitions. 50,000 scans were summed and 2.5 Hz line broadening applied for plotting. D_2O (50%, Chellex treated to remove paramagnetic ions) was added as an internal lock. These experiments typically required 0.8 µmol sample.

Nucleoside Analyses. Nucleoside analyses were performed on a Spectra-Physics HPLC in the laboratory of Dr. L. Brox. Results were calibrated with standards of accurately known concentration.

Scintillation Counting. Aqueous samples were counted in a aquasol (NEN) usually 1 ml sample and 9 ml fluor. Samples adsorbed onto paper were first eluted with water. Acid precipitable counts were determined by pipetting small aliquots onto Whatman GF/C glass fibre filters (presoaked with ice-cold 5% trichloroacetic acid) mounted on

sintered glass supports. Filters were then washed with 5% TCA (5-10 ml) and then dried with ethanol (5 ml) and air. Non-specific adsorbtion was minimized by adding 10 mM inorganic pyrophosphate to TCA solutions. Samples were counted after adding 5 ml fluor (8g Scintimix 3, Koch-Light Laboratories, per 2 L toluene). Overlap corrections were applied where necessary.

Sedimentation Velocity Measurements. Alkaline sedimentation velocity experiments were performed in a Beckman Model E analytical ultracentrifuge in Vinograd Type I Epon cells at 56000 rpm, 25° in an AnF Ti (four hole) rotor. Cell paths were scanned at 8 minute intervals using a multiplexer accessory. Samples (20 µl 2.5 A_{260} DNA + 5 µl 1.0 M NaOH) were first heated 100°, 2' and layered during centrifugation onto 0.55 ml 0.9M NaCl, 0.1 M NaOH. Calculations were according to Studier (1965), the correction of S to $S^{\circ}_{20,w}$ being $S^{\circ}_{20,w} = 1.16 \times S$ in this buffer. Weight average molecular weights were calculated by

$$M_w = \left(\frac{S^{\circ}_{20,w}}{0.0528}\right)^{2.5}$$

this equation is accurate to circa 4S and is derived from natural DNA's (Hirose et al, 1973).

Ultraviolet Dosimetry. U.V. dosimetry was determined using potassium ferrioxalate actinometry (Jager, 1967). This technique involves the preparation of a potassium ferrioxalate [$K_3\ Fe(C_2O_4)_3 \cdot 3H_2O$] solution which is irradiated and the reduced iron measured spectrophotometrically using a phenanthroline indicator. Quantum yield was taken to be 1.26 mol/einstein and actinometer absorbtion 100% at 254 nm (Jager, 1967).

Preparation of Double Stranded Synthetic DNA's

Reaction Conditions. Polymerase reactions varied slightly in composition typically these contained 50 mM KPi pH 7.5, 12 mM $MgCl_2$, 0.02% NaN_3, 1 mg/ml autoclaved gelatin, 6 mM total dXTP's (in a ratio appropriate to the polymer being replicated), 0.1-0.5 A_{260} primer DNA and DNA polymerase I (Morgan et al, 1974). DNA's containing unusual bases were prepared by substituting the appropriate deoxynucleoside triphosphate. In order to inhibit poly d(AT) synthesis (see Chapter III) netropsin (1 µg/ml) or TANDEM (2 µg/ml) were added when necessary. Also, because the processivity of DNA polymerase I on poly d(AT) is markedly ionic strength dependent (Bambara et al, 1978), the concentration of KPi was increased to 125 mM in later syntheses.

Polymerase reactions were followed using the ethidium bromide fluorescence assay described above. Small aliquots were added to 2 ml pH 8 ethidium solution and the fluorescence measured relative to standard DNA's of known concentration. Contamination of poly pyrimidine·polypurine DNA's by poly d(AT) could be measured by heat denaturing, cooling and remeasuring the fluorescence. Poly d(AT) - because of its self-complementary nature - "snaps back". Reactions were stopped before degradation started to occur.

Purification. Reactions were stopped by adding EDTA to 20 mM and proteinase K to about 50 µg/ml. Samples were concentrated with air at 37°, passaged on Biogel and the excluded material retained. Occasionally laminar, partly birefingent gels formed during concentration. These were redissolved by heating at 45-60°. Recoveries of purified DNA's typically averaged around 80%.

Characterization. Table 2-1 summarizes some of the properties

Table 2-1

Physical properties of synthetic pyrimidine·purine DNA's.

DNA	Tm[1](C°)	$A_{260/280}$	Specific[2] fluorescence	ε_{260}
poly dA·poly dT	49.5	—	--	6000[3]
poly d(TTC)·poly d(GAA)	58.5 ± 1.2	1.78 ± 0.03	64.5 ± 5.2	6300[4]
poly d(TC)·poly d(GA)	64.2 ± 1.2	1.83 ± 0.04	65.2 ± 7.3	6500[4]
poly d(TCC)·poly d(GGA)	71.8 ± 2.1	1.88 ± 0.03	69.2 ± 8.6	6800[4]
poly dG·poly dC	86	--	--	7100[3]

1. T_m's are in 10 mM Tris·HCl pH 8.0, 5 mM NaCl, 0.1 mM EDTA.

2. Fluorescence of 10 µl 1.0 A_{260} DNA with a calf thymus DNA standard reading 70.

3. Wells et al, 1970.

4. Murray, 1972.

of pyrimidine·purine DNA's prepared over several years. Extinction coefficients are added for reference purposes. DNA preparations were always checked by T_m measurement for purity and identity. Other checks of identity (not always performed) included nearest neighbour analysis of RNA transcripts, base composition, buoyant densities and transcription in the presence of a limited number of rNTP's (eg. UTP and CTP for poly d(TC)·poly d(GA)).

Molecular weights were determined by several methods. Agarose gel electrophoresis showed broad molecular weight distributions from 0.5 -25 Kb. Electron microscopy showed these DNA's to be simple unbranched linear molecules which could, if desired, be separated into discrete size classes by chromatography on Biogel A-150 m (50-100 mesh). Alkaline sedimentation velocity showed these DNA's to be composed of single strands 0.2-0.6 Kb in length. Kinase endlabelling of phosphatase treated DNA's gave anomalously high molecular weights (2-3 Kb) presumably due to failure to label at internal nicks (Masamune et al, 1971).

Preparations of Single Stranded Synthetic DNA's

Alkaline Buoyant Density. The rather low molecular weights of these synthetic DNA's made strand separation on the basis of differential buoyant densities difficult. The best separation achieved is shown in Figure 2-1 this was only made possible by first fractionating denatured DNA on Biogel A 150m agarose in 5 mM NaOH, 0.1 mM EDTA and banding the excluded material. This technique rarely produced DNA's of the requisite purity by itself so crude DNA's from density gradients were normally further purified as described below.

Polypyrimidine DNA's by Depurination. DNA's were incubated with 67% formic acid at 37° for 16-20 hours (Harwood and Wells, 1970).

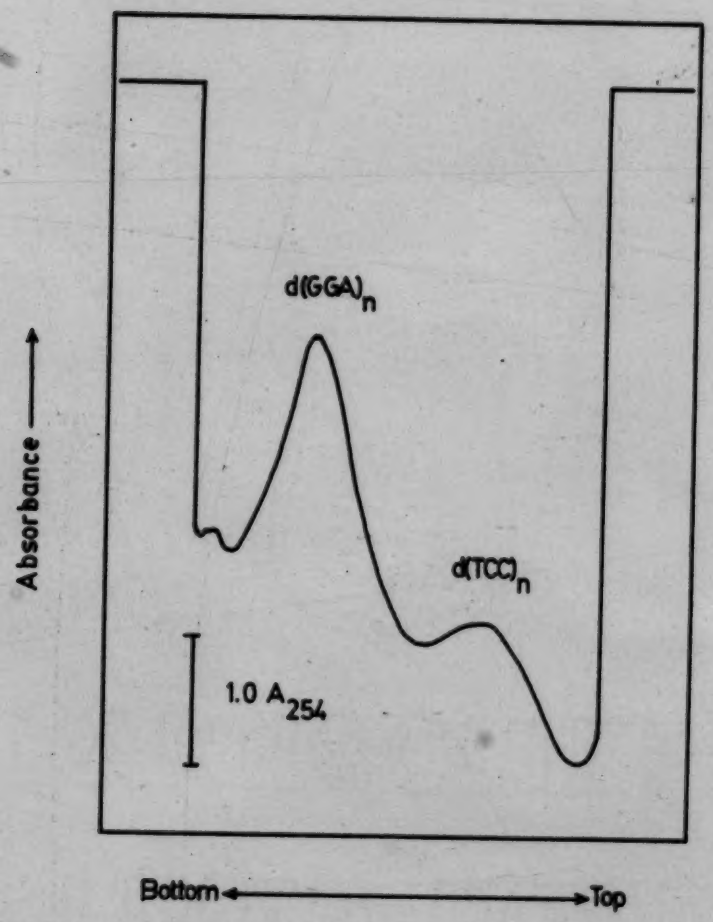


Figure 2-1 Separation of poly d(TCC) from poly d(GGA) by alkaline CsCl centrifugation. Conditions as per Morgan et al (1974), initial $\rho = 1.805$ g/ml by refractometry. Tubes were punctured and the absorbtion profile recorded with a Pharmacia column monitor at 254 nm.

The formic acid was removed by lyophilization and the residue resuspended in 0.2 M NaOH and incubated at 37° for 5 hours. The pH was then adjusted to pH 11 (pH paper) and the product chromatographed on Biogel. Yields were good (typically 65-95%) and completely free of contaminating polypurine DNA. DNA's prepared in this way had molecular weights (by alkaline sedimentation velocity) indicating polymers of 40-70 nucleotides long suggesting that some depyrimidation also occurs during acid treatment. These are only rough estimates since the equations relating $S^{\circ}_{20,w}$ and M aren't calibrated in this range.

For nucleoside analyses samples were digested to completion with a mixture of micrococcal nuclease, snake venom phosphodiesterase* and bacterial alkaline phosphatase. To 100 µl of digest (containing 0.3-0.5 nmol of each nucleoside) was then added 4.5 µl 9.2 M perchloric acid (on ice) and left 15'. After centrifuging (clinical centrifuge) 100 µl alamine-freon (6.2:18.8 V/V) was added vortexed and recentrifuged. 80 µl of the top phase (there are three) was removed, neutralized with 4 µl 0.1 M NaOH and stored frozen for analysis (the later steps are a modification of a procedure used in Dr. Brox's laboratory). Nucleoside analyses normally agreed within ±2% of the expected pyrimidine base ratios but tended to under-recover purine nucleosides. In some cases a small peak of UdR was observed. This may indicate the occurance of deamination during polypyrimidine preparation or possibly the incorporation of dUTP during duplex synthesis.

Product characteristics are indicated in Table 2-2. For reference poly dT and poly dC are included.

* One preparation of SVPD (Boehringer-Mannheim) contained a "cytidine deaminase" activity. Substitution of PL Biochemicals SVPD avoided this problem.

Table 2-2

Properties of polypyrimidine DNA's.

DNA	λ_{max}	$E_{\lambda max}$	E_{260}	$A_{260/280}$
poly dT	264	8520	8100	1.52
poly d(TTC)	267	8800	8020	1.38 ± 0.03
poly d(TC)	268	9100	8650	1.34 ± 0.03
poly d(TCC)	268	10400	9250	1.27 ± 0.04
poly dC	268	7400	--	0.81

Poly dT and poly dC data are from Ts'o et al (1966). Other polymer extinction coefficients were determined using the method of continuous variations (Lee et al, 1979) and are averages of data obtained by myself and Dr. J.S. Lee. Buffer is TE.

Polypurine DNA's. Polypurine DNA's were prepared using a procedure of Dr. J.S. Lee's modified to improve purity and yield. Polypyrimidine strands were selectively degraded under the conditions indicated in Figure 2-2 and purified in alkali on Biogel. Molecular weights (alkali sedimentation velocity) were similar to polypyrimidine DNA's. The preparation of polypurine DNA's by exonuclease III presumably depends upon both the tendency of polypurine DNA's to form ordered complexes (Chapter IV) and a slight intrinsic specificity for pyrimidines*. Yields averaged 30-50% the best yields being found in poly d(GA) and poly d(GGA) preparations. Poly d(GA) and poly d(GGA) were typically contaminated with 4-5% pyrimidine and poly d(GAA) about 10%. Purity could be improved by purification and redigestion but low yields (primarily column losses) made this impractical. Free polypurine strand concentrations were determined by titration with the appropriate polypyrimidine strand of known concentration because polypurine extinction coefficients tended to vary rather irreproducibly. Strand identities were easily checked by comparing their T_m with those reported earlier (Lee et al, 1980 and Chapter IV).

Preparation of Polypurine RNA's

Polypurine RNA's were prepared by RNA polymerase transcription of the appropriate DNA's as described by Morgan (1970). Reactions contained 200 mM NaCl, 10 mM $MgCl_2$, 50 mM Tris·HCl pH 8.0, 0.1 mM cacodylate, 2.0 mM DTT, 6% v/v glycerol, 0.1 mg/ml autoclaved gelatin, 1.0mM ATP, 1.0 mM GTP and template (~200 μM). After synthesis stopped RNA and

* No evidence of exonuclease III specificity has previously been reported although some of Wu's data (eg. Ghangas et al, 1973) could conceivably be interpreted as sequence dependent phenomena.

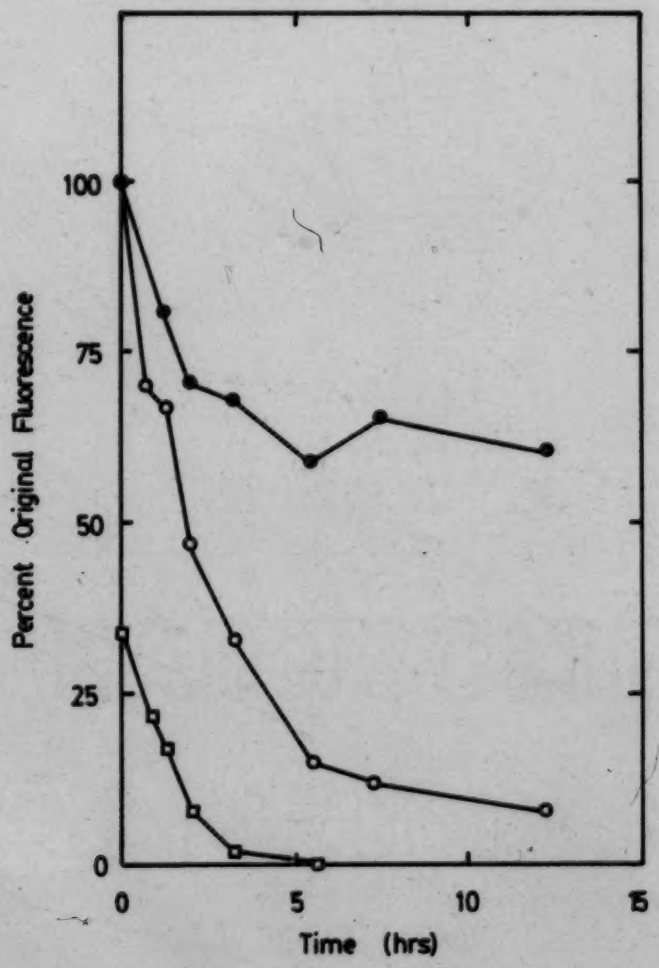


Figure 2-2 Preparation of poly d(GAA) by exonuclease III digestion. The reaction contained 8 mM $MgCl_2$ 40 mM KPi pH 6.4, 4 mM DTT, 100 mM NaCl and exonuclease III (100 U). Initial DNA concentration was 20 A_{260} since high DNA levels were found to improve yields. The reaction was incubated at 37° and samples either added directly to 2 ml pH 8 ethidium buffer (-o-) or first added to an excess of poly d(TTC) (-●-). Note the complete degradation of a poly d(AT) contaminant as followed by a decrease in heat resistant fluorescence (-□-).

DNA were isolated off Biogel in TE and $MgCl_2$(4 mM), $CaCl_2$(0.5 mM and DNase I (50 U/ml) added. Reactions were digested, heat denatured (100°, 2'), more DNase I added and digested a second time (4 hours to overnight) see Figure 2-3. RNA's prepared in this manner were typically contaminated with 2-5% polypurine DNA's. Nearest neighbour and nucleoside analyses showed the expected sequences and compositions. Other aspects of RNA synthesis are discussed in more detail in Chapter IV.

Preparation of [^{3}H] Labelled T7 DNA

^{3}H labelled T7 DNA was prepared according to the procedure of Kelly and Thomas (1969) on E.coli CR34 (ton A, λ^s, thy^-, thr^-, leu^-) in M63 media supplemented with 2 mg/ml glucose, 40 µg/ml thymine. Prior to addition of T7 innoculum bacteria were labelled with 50 µCi TdR. After extraction with phenol phage DNA was purified on Biogel A 150 m and had a final specific activity of 3800 cpm/µmol Pi. Agarose gels showed only one band of apparent molecular weight 25×10^6.

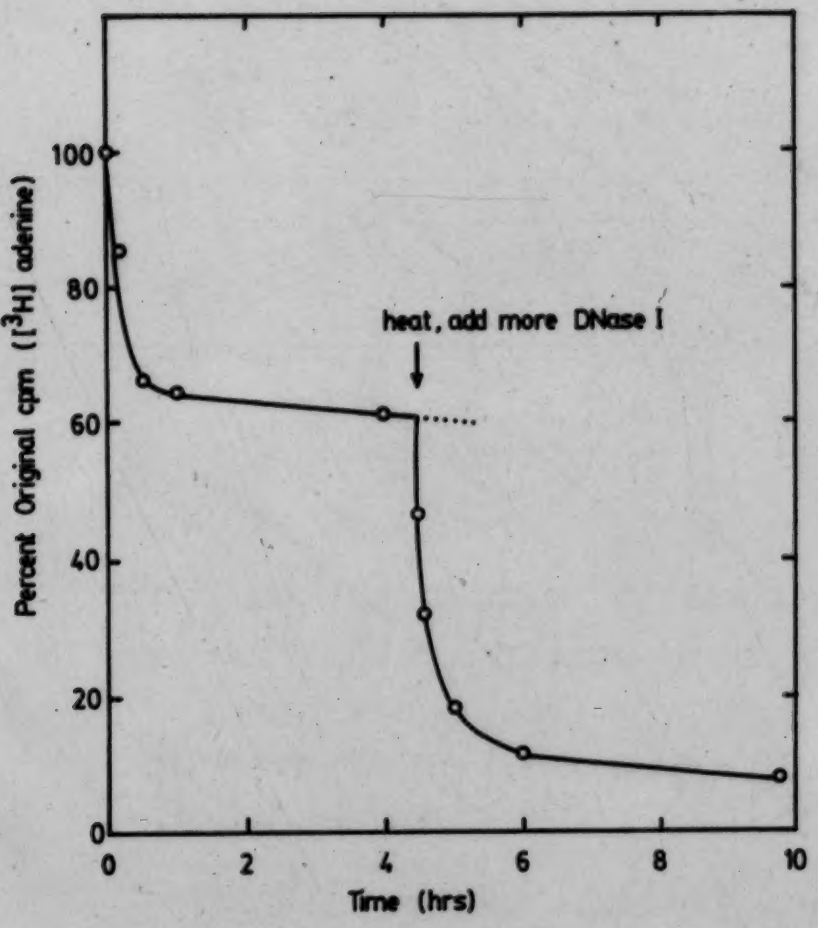


Figure 2-3 DNase I digestion of poly d(GAA)·poly d(TTC), poly r(GAA) mixture. The reaction was as described above and contained 180 mM DNA and (^{3}H-labelled with dATP) and approximately 200 mM unlabelled. After 4.5 hours DNase digestion 100 U/ml more DNase I was added as indicated. Notice the resistance of this material to nuclease digestion. 100% corresponds to 3000 cpm.

III. Specific Inhibitors of DNA Polymerase I catalysed de novo DNA Synthesis

Introduction

De Novo Synthesis

E. coli DNA polymerase I can, in the absence of any primer DNA, synthesize de novo a number of repeating sequence synthetic DNA's. After a lag period in the presence of two nucleoside triphosphates and Mg^{+2} the synthesis of poly d(AT) (Schachman et al, 1960), poly dG·poly dC (Radding et al, 1962), poly dI·poly dC (Inman and Baldwin, 1964) and poly d(IC) (Burd and Wells, 1970) has been observed. In the presence of all four triphosphates the synthesis of a polymer identified only as poly d(T,G)·poly d(C,A) has also been observed to appear in reactions primed with poly d(TTG)·poly d(CAA) (Coulter, 1974). Of these and other synthetic polymers poly d(AT) appears to be by far the preferred template of DNA polymerase I.

The origin of de novo synthesis is the subject of some debate. One hypothesis is that polymerase I is capable of synthesizing small oligonucleotides of random sequence. Those oligonucleotides capable of being copied by a slippage mechanism (homo and simple repeating sequence oligonucleotides) are then selectively replicated (Kornberg et al, 1964). A second, and more likely hypothesis, is that DNA polymerase preparations are always contaminated by trace levels of oligonucleotides which can serve as primers for polymer synthesis (Kornberg, 1980). Experimentally it is observed that de novo synthesis is sensitive to reaction conditions and polymerase purity (Burd and Wells, 1970); Morgan et al, 1974), the 5' exonuclease activity of DNA polymerase I (Brutlag and Kornberg, 1970)

and the presence of DNA binding compounds (Olson et al, 1972; Wartel et al, 1974).

Regardless of the origin of de novo synthesis it is a significant problem associated with the replication of synthetic polynucleotides. In particular poly d(AT) often arises de novo in reactions primed with other DNA's and, left to itself, can completely divert synthesis away from the more desired polymer. A number of procedures have been devised in an attempt to selectively inhibit poly d(AT) synthesis in primed reactions. High deoxyribonucleoside triphosphate levels (Morgan et al, 1974), high ionic strengths and high initial primer DNA to polymerase ratios are all helpful. More specifically Wells and coworkers have suggested the use of the small DNA binding oligopeptide antibiotic netropsin (Berman et al, 1979; Wartell, Larson and Wells, 1974) to selectively inhibit the synthesis of A·T rich DNA's such as poly d(AT). Unfortunately the presence of netropsin in polymerase reactions was found to have a rather undesirable side effect as described below.

De Novo Synthesis of poly d(TC)·poly d(GA)

Because poly d(AT) kept appearing during attempts to synthesize poly d(TTC)·poly d(GAA) I added netropsin at 1 μg/ml in an attempt to specifically inhibit its synthesis. The product of these reactions turned out to have a T_m consistant with it being poly d(TC)·poly d(GA) but 8° higher than that expected of poly d(TTC)·poly d(GAA). Figure 3-1(a) illustrates that this polymer could, in fact, be synthesized in the absence of primer DNA with kinetics typical of de novo synthesis. The T_m of this material in TE + 5mM NaCl was 63°, Figure 3-1(b).

The identity of this polymer was established as follows. Because polypyrimidine-polypurine DNA's have been shown to form triplexes

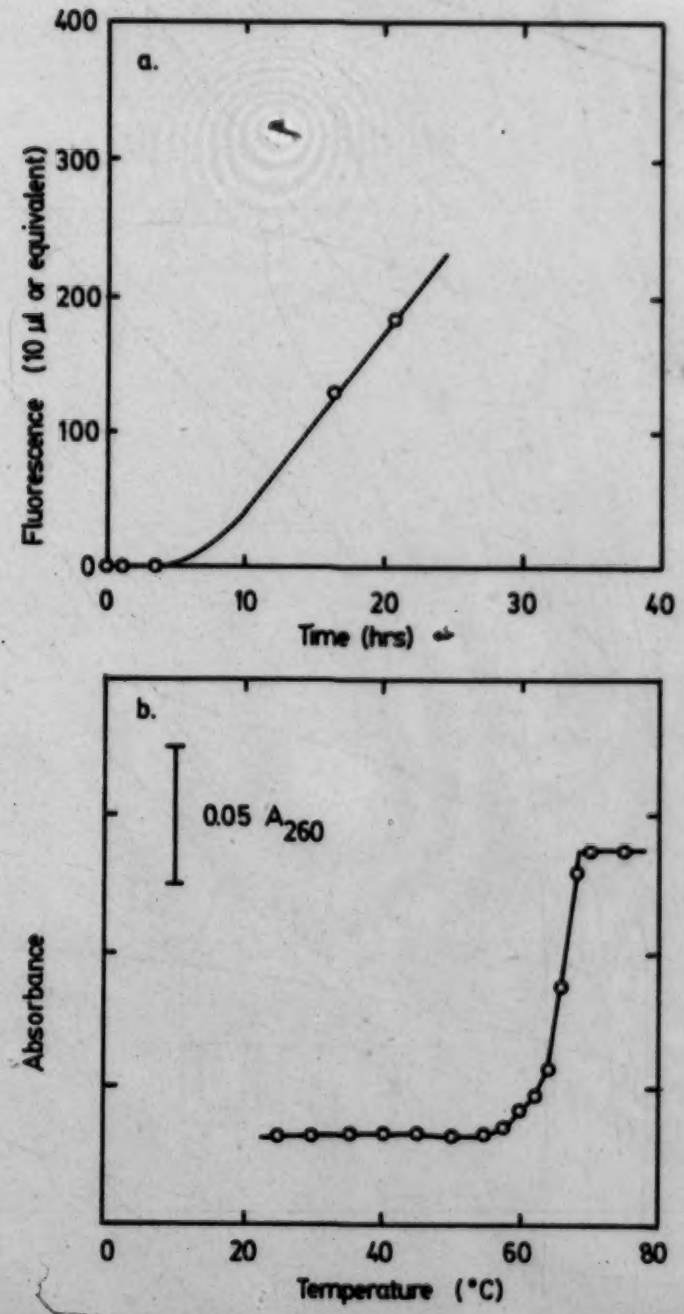


Figure 3-1 (a) De novo synthesis by E.coli polymerase I. Normal polymerase conditions 1.3 mM each dATP and dTTP, 0.67 mM each dGTP and dCTP, 1 µg/ml netropsin. No after-heat fluorescence. (b) T_m of DNA synthesized de novo in the presence of netropsin.

at low pH's (Lee et al, 1979) I incubated this material at pH 5.0. This resulted in a loss of fluorescence suggesting formation of a new structure. This rearrangement is typical of that which accompanies triplex formation (non-pyrimidine purine DNA's are stable under these conditions). As expected digestion with the single-strand specific endonuclease S1 destroyed one-half the total duplex fluorescence originally present. Finally nearest-neighbour analysis of an RNA transcript synthesized in the presence of [α]-^{32}P labelled ATP showed >95% of label transferring to GMP.

These results were obtained with one particular preparation of DNA polymerase but at least two other preparations were also observed to synthesize, on occasion, a polymer with a 63° T_m.

These data suggest that in addition to the aforementioned polymers DNA polymerase I can also synthesize de novo poly d(TC)·poly d(GA) under conditions where the synthesis of poly d(AT) is suppressed. This led me to try two more series of experiments. Firstly to find a rational for why netropsin allows poly d(TC)·poly d(GA) synthesis to take over in reactions primed with poly d(TTC)·poly d(GGA) and secondly to look for alternate inhibitors of poly d(AT) synthesis.

Netropsin Binding to poly d(TTC)·poly d(GAA)

Netropsin has been shown to bind to synthetic polymers such as poly d(AT) and poly dA·poly dT with apparent binding constants of circa 1×10^6 and polymers such as poly dG·poly dC about 300 fold less. The physical binding site size (assuming neighbour exclusion) being ~3 base-pairs (Wartell, Larson and Wells, 1974). Binding has been postulated to involve hydrogen-bonding between three netropsin amide hydrogens and thymine O2 and/or adenine N3 positions (Berman et al, 1979). The presence

of a guanine N2 amino group is hypothesized to block this. Presumably the reason why netropsin allows poly d(TC)·poly d(GA) synthesis to take over poly d(TTC)·poly d(GAA) primed reactions is simply that its specificity is such that it can bind poly d(TTC)·poly d(GAA) but not poly d(TC)·poly d(GA). The effect is to cause selective replication of the latter polymer.

Netropsin does indeed bind poly d(TTC)·poly d(GAA) and not poly d(TC)·poly d(GA) as shown in Figure 3-2. In CsCl 5 µg/ml netropsin shifted the buoyant density of poly d(TTC)·poly d(GAA) by 38 mg/ml (1.686 → 1.648) while having essentially no significant effect on poly d(TC)·poly d(GA) (1.711 → 1.709) density. These results are of some practical value and also suggest that previous models of netropsin-DNA complexes should be modified to allow for the presence of a G·C base-pair within the binding site. CPK models show that a hydrogen-bonded complex between netropsin and an ApApG sequence can be built but whether such a structure really exists is open to question.

TANDEM - A Selective Inhibitor of poly d(AT) Synthesis

Netropsin is a good inhibitor of poly d(AT) synthesis however it is not specific enough to be of any use when replicating polymers such as poly dA·poly dT or poly d(TTC)·poly d(GAA). This led me to investigate the properties of TANDEM in the hope that it would be a more suitable inhibitor of poly d(AT) replication.

TANDEM

TANDEM (des-N-tetramethyl-triostin A) is a chemically synthesized analogue of the naturally occurring quinoxaline antibiotic triostin A. It consists of a disulfide cross-bridged, cyclic octapeptide dilactone with two attached quinoxaline-2-carbonyl ligands (Ciardelli *et al*,

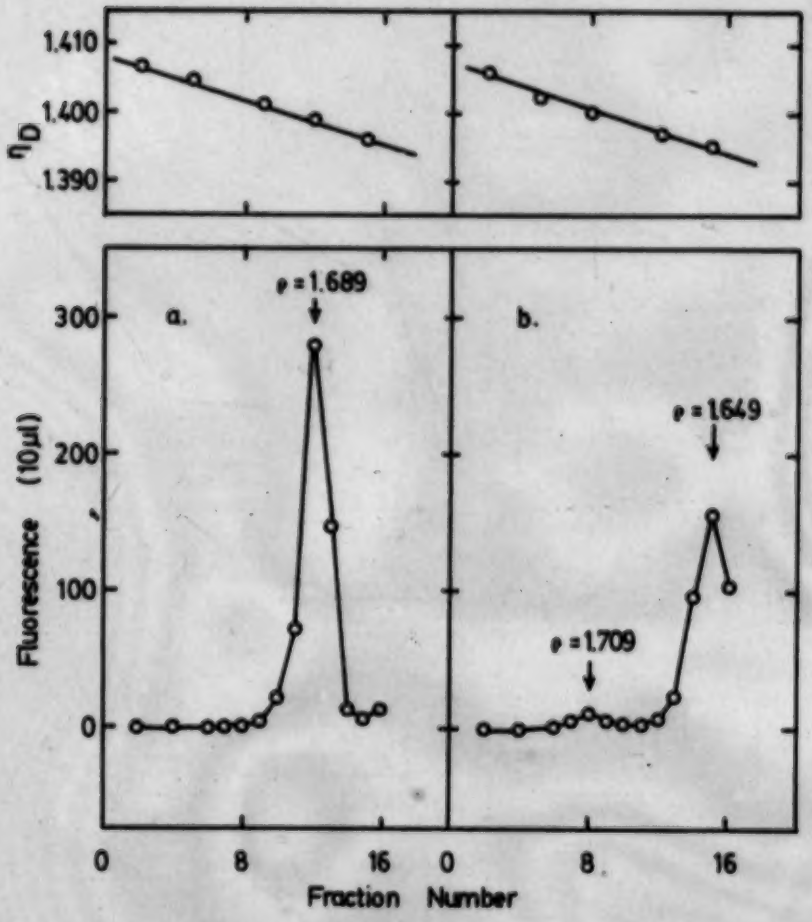


Figure 3-2 Netropsin binding to poly d(TTC)·poly d(GAA) as demonstrated by buoyant density shifts. Gradients contained DNA (3% poly d(TC)·poly d(GA) 97% poly d(TTC)·poly d(GAA) - presized on Biogel A 150m), 10 mM Tris·HCl pH8, 0.1 mM EDTA, CsCl (initial density 1.707 g/ml) either without (a) or with (b) 5 µg/ml netropsin. Centifugation (at 8°, for 84 hours) and fractionation were as described earlier. Reference buoyant densities are 1.686 and 1.711 for poly d(TTC)·poly d(GAA) and poly d(TC)·poly d(GA) respectively (Wells et al, 1970).

1979) (Figure 3-3). Binding to DNA involves bifunctional intercalation and shows an extremely high affinity for poly d(AT) with an intrinsic binding constant $K(0) = 1.50 \times 10^{7} M^{-1}$ (Lee and Waring, 1978). Binding is cooperative and highly specific suggesting that TANDEM should be a suitable inhibitor of poly d(AT) synthesis.

Effect of TANDEM on poly d(AT) Synthesis

Figure 3-4 shows the effect of TANDEM on poly d(AT) synthesis. Complete inhibition of synthesis is seen at drug to phosphate (D/P) levels of 0.5 while lower D/P ratios are less inhibitory. Since approximately 98.8% of the TANDEM binding sites should be occupied under the conditions of Figure 3-4 when D/P = 0.5 (McGhee and von Hippel, 1974; erratum, 1976) this suggests that inhibition of polymerization is simply due to TANDEM blocking polymerase access to the template. At lower concentrations of TANDEM not all sites can be occupied and this would allow polymerization to take place. The kinetics observed at D/P = 0.1 indicate that slow polymerization eventually titrates out the drug causing an exponential increase in the rate of synthesis.

TANDEM Specificity

Table 3-1 shows the effect of TANDEM on the rate of synthesis of a variety of different synthetic DNA's. Clearly, TANDEM has no significant inhibitory effect on the rate of synthesis of these DNA's except for poly d(TAC)·poly d(GTA). In order to demonstrate that inhibition of synthesis (or lack of it) correlates with drug binding the effect of TANDEM on the melting temperature of these polymers was examined. TANDEM has no effect on the T_m of poly d(TTC)·poly d(GAA) for example whereas as is shown in Figure 3-5 it can be seen that TANDEM increases the T_m of poly d(TAC)·poly d(GTA) by nearly 6°. The T_m changes for these and

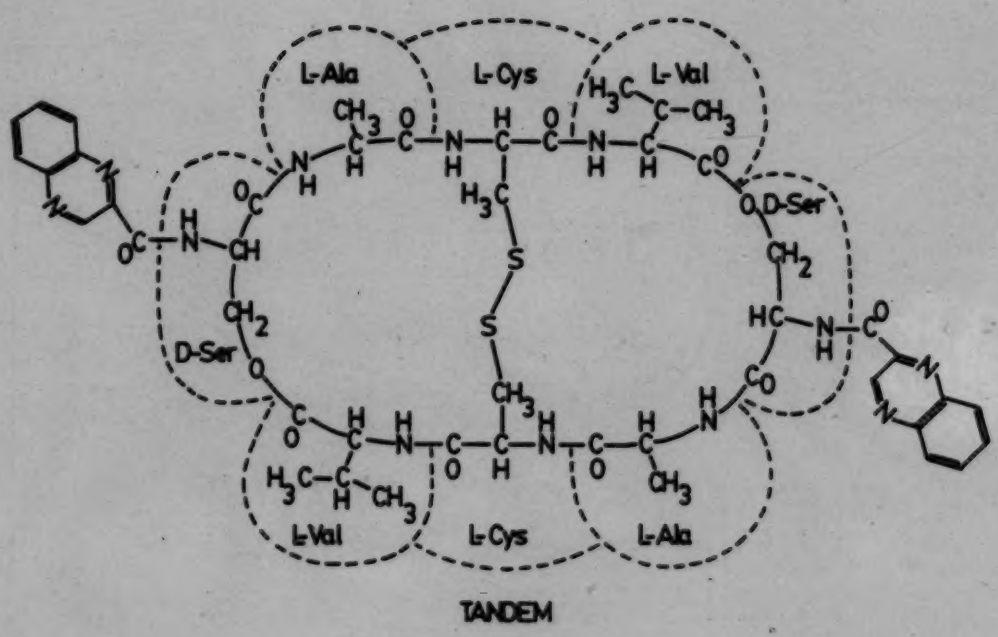


Figure 3-3 The structure of des-N-tetramethyl triostin A (TANDEM) (Lee and Waring, 1978).

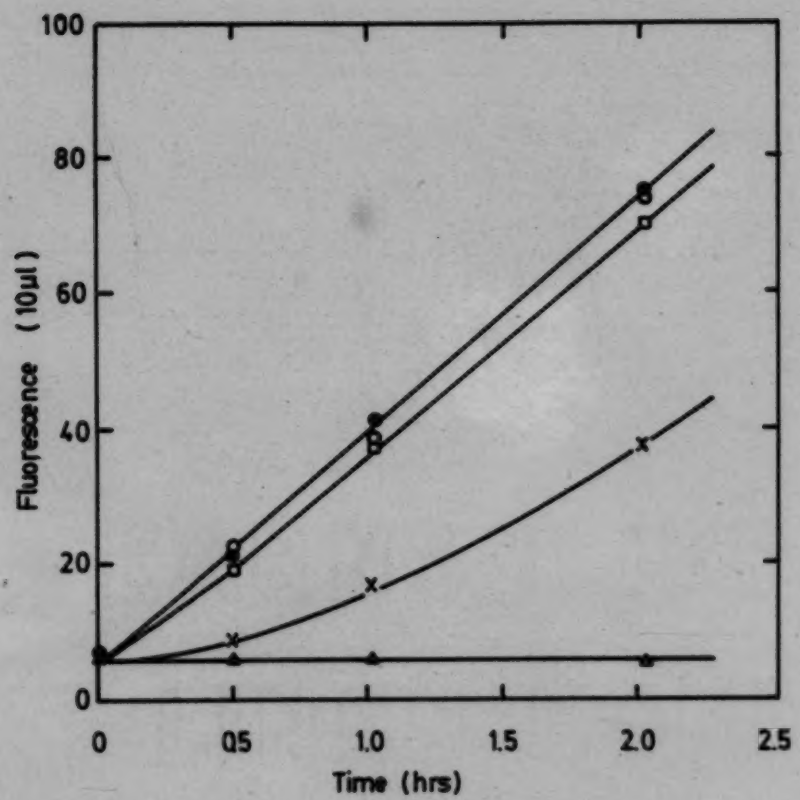


Figure 3-4 Effect of TANDEM on poly d(AT) synthesis. DNA polymerase reactions (50 µl) contained 0.14 U E.coli DNA polymerase I, 1.5 mM each of dATP, dCTP, dGTP and dTTP, 10 µm poly d(AT) and TANDEM at 5 µM (-Δ-), 1.0 µM (-X-) or 0.1 µM (-□-). Two controls are shown, one contains 1.0% ethanol since this was the TANDEM solvent; (-o-) - ethanol, (-●-) + ethanol.

Table 3-1

The effect of TANDEM on the Rate of Synthesis of various synthetic DNA's.

DNA	Initial Rate as a Percent of Control	D/P Ratio
Poly d(AT)	8.8	0.10
poly dA·poly dT	121	0.13
poly d(TTC)·poly d(GAA)	102	0.13
poly d(TC)·poly d(GA)	95	0.09
poly d(TG)·poly d(CA)	99	0.26
poly d(TAC)·poly d(GTA)	7.9	0.21
poly d(ATC)·poly d(GAT)	101	0.20

Rates were determined as in Figure 3-4 using 5 μM TANDEM and 25-55 μM primer DNA.

Error is approximately ±5%.

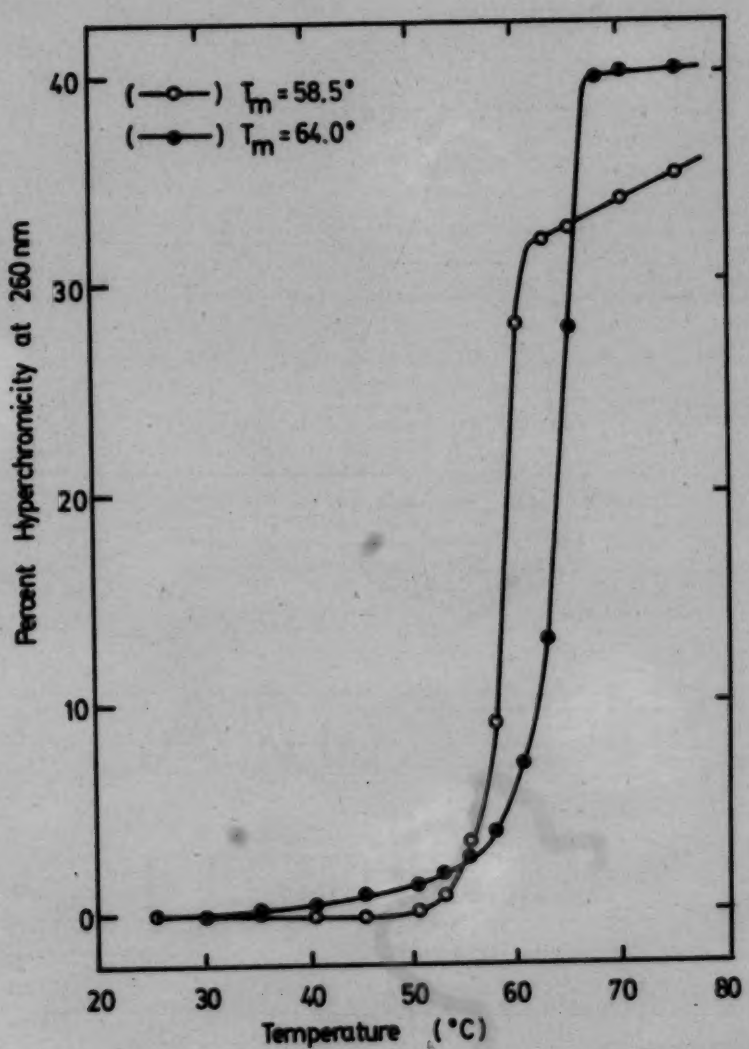


Figure 3-5 Effect of TANDEM on poly d(TAC)·poly d(GTA) melting. Sample contained 60 μM DNA, TE, 5mM NaCl and ethanol (0.3%) (-o-) or ethanol plus TANDEM (5 μM)(-●-).

other polymers in the presence of TANDEM are summarized in Table 3-2. Comparison with Table 3-1 shows that there is a strong correlation between drug binding and inhibition of synthesis. Indeed it would appear that poly d(TAC)·poly d(GTA) is the only polymer for which TANDEM cannot be used to prevent poly d(AT) synthesis. Figure 3-6 shows an example of how these observations can be of practical value. Since poly d(TC)·poly d(GA) did not bind TANDEM (Tables 3-1 and 3-2) it could be selectively replicated in a reaction mixture originally primed with both poly d(TC)·poly d(GA) and poly d(AT). This technique has allowed us to continue replicating DNA's which previously could not be further replicated because of poly d(AT) contamination.

TANDEM Inhibition of de novo Synthesis in Unprimed Reactions

TANDEM inhibits the synthesis of poly d(AT) in reactions primed with poly d(AT). Can it also be used to inhibit poly d(AT) synthesis in unprimed reactions and if so does this lead to the de novo synthesis of other polymers? Figure 3-7 illustrates the effect of TANDEM on de novo synthesis in reactions catalyzed by E.coli DNA polymerase. As expected, in the absence of TANDEM the E.coli polymerase synthesized poly d(AT) after a short lag period. In the presence of 5.0 μM TANDEM this reaction was completely inhibited and after a much longer lag period poly dA·poly dT synthesis eventually took place. This result is unusual since poly dA·poly dT synthesis is normally very difficult to detect (Burd and Wells, 1970) presumably because of the E.coli polymerase's preference for producing poly d(AT). Figure 3-7 shows clearly that the E.coli polymerase can synthesis poly dA·poly dT de novo as long as a method is employed to inhibit poly d(AT) replication.

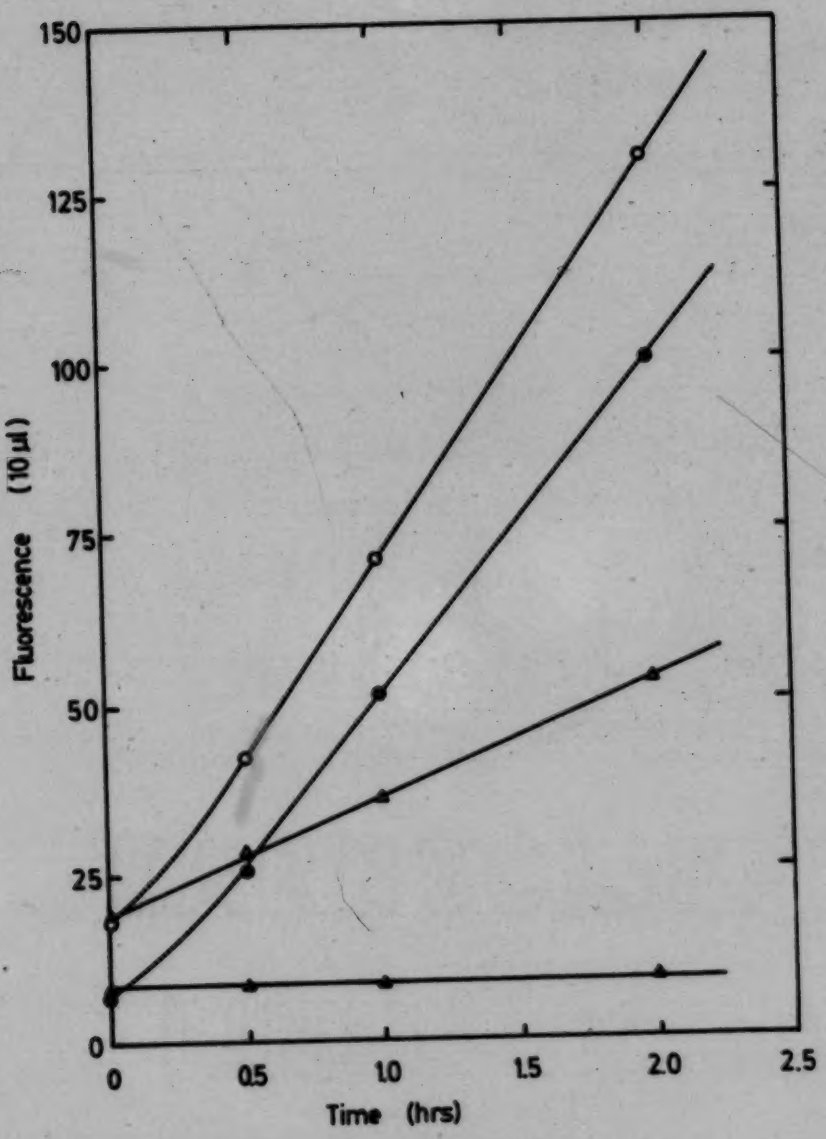


Figure 3-6. Selective inhibition of poly d(AT) synthesis in the presence of poly d(TC)·poly d(GA) synthesis and TANDEM. DNA polymerase reactions were identical to those in Figure 3-4 except 28 µM poly d(TC)·poly d(GA) was added in addition to poly d(AT). Control (no TANDEM) before heat (-o-), after heat (-●-) plus 5.0 µM TANDEM before heat (-Δ-), after heat (-▲-). After heat fluorescence represents fluorescence due to poly d(AT) alone.

Table 3-2

The effect of TANDEM on the T_m of various synthetic DNA's.

DNA	ΔT_m (°C)	D/P Ratio
poly d(AT)	5.5	0.018
poly d(AT)	8.5	0.050
poly dA·poly dT	0.5	0.096
poly d(TTC)·poly d(GAA)	0*	0.053
poly d(TTG)·poly d(CAA)	0*	0.063
poly d(TC)·poly d(GA)	0	0.098
poly d(TG)·poly d(CA)	0	0.096
poly d(TAC)·poly d(GTA)	5.5	0.083
poly d(ATC)·poly d(GAT)	0	0.110

T_m's were determined as in Figure 3-5 ΔT_m is the difference between T_m (+ TANDEM) minus T_m (-TANDEM). Poly d(AT) data is taken from Lee and Waring (1978).

* Determined by Dr. J.S. Lee.

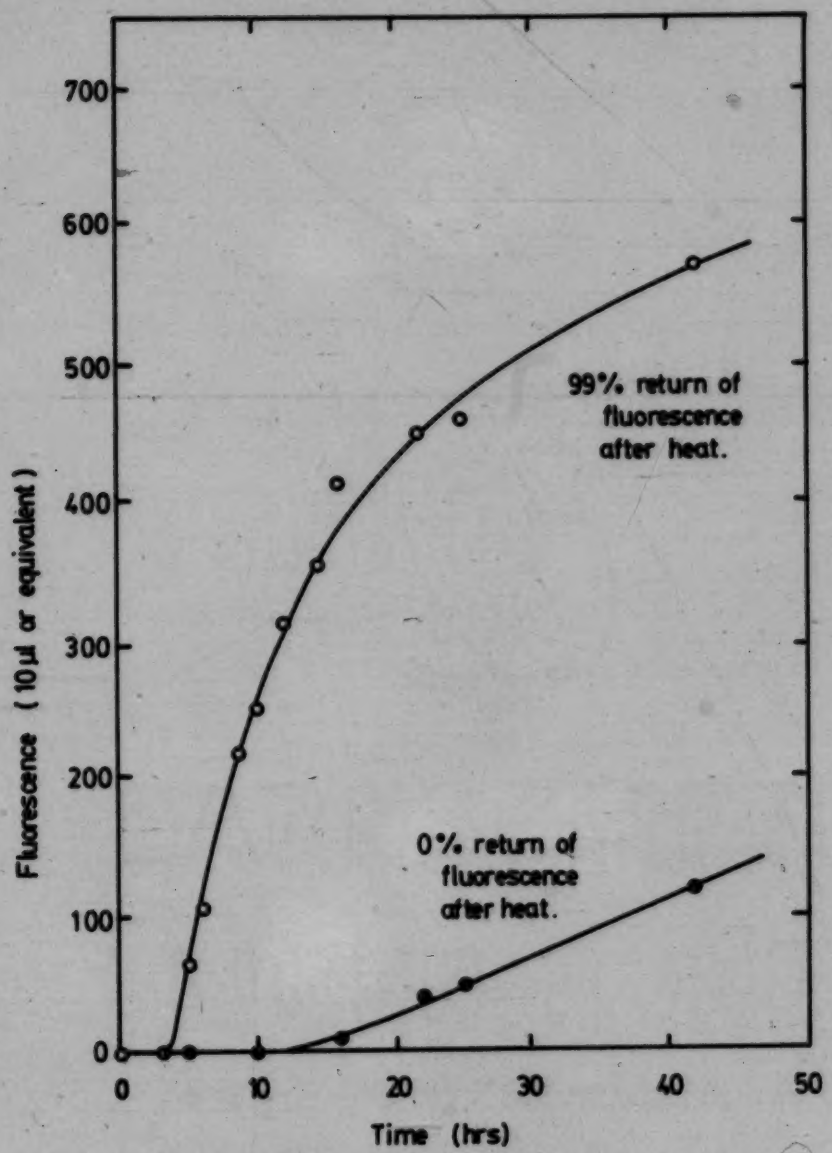


Figure 3-7 Effect of TANDEM on unprimed poly d(AT) synthesis. Polymerase reactions contained 0.72U of polymerase, 3mM dATP, 3mM dTTP and salts as described previously. Control (no TANDEM) (-o-), plus 5 µM TANDEM (-●-).

Discussion - TANDEM Binding Specificity

TANDEM binding specificity can be exploited to inhibit the synthesis of poly d(AT) in reactions primed with a large variety of other synthetic DNA's. This specificity is great enough that even DNA's with high A·T contents can be replicated in the presence of TANDEM. These observations are of considerable practical value since this procedure eliminates a major problem associated with the preparation of synthetic repeating sequence DNA's.

Of more theoretical interest is the observation that TANDEM inhibits the replication of poly d(TAC)·poly d(GTA) while having no effect on the replication of poly d(ATC)·poly d(GAT). Although it is difficult to deduce binding constants from changes in T_m (McGhee, 1976) the observed $\Delta T_m = 5.5°$ at D/P = 0.083 for poly d(TAC)·poly d(GTA) can be interpreted to mean that the binding of TANDEM to this polymer is probably weaker than to poly d(AT) ($\Delta T_m = 5.5°$ at D/P = 0.018 (Lee and Waring, 1978). Similarly the fact that for poly d(ATC)·poly d(GAT) no change in the T_m could be detected suggests that the binding constant of TANDEM is less than 10^3 M^{-1} (Lee and Waring, 1978). Curiously a number of different antibiotics unrelated to the quinoxalines show a similar pattern of selectivity. The anthracycline antibiotics nogalamycin and ruticulomysin also appear to be able to discriminate between poly dA·poly dT and poly d(AT) binding to the latter but not the former (Olson et al, 1972). Furthermore actinomycin D also shows this selectivity binding to poly d(TAC)·poly d(GTA) but not to poly d(ATC)·poly d(GAT) (Wells and Larson, 1970).

It is difficult to rationalize this aspect of TANDEM specificity without detailed binding data; the effect, for example of G·C base-

pair orientation on TANDEM binding is impossible to predict. However it is tempting to speculate that TANDEM actually recognizes a TpA rather than ApT sequence. This is not unreasonable since in poly d(AT) these sequences have been proposed to have quite different conformations (Klug et al, 1979) and differences like these can be expected to affect drug binding. Indeed if one considers that the poly d(AT) binding site size is approximately eight nucleotides it is conceivable that TANDEM actually recognizes only one of the two dinucleotide units present in poly d(AT). A recent crystal structure of TANDEM (Viswamitra et al, 1981) suggests that hydrogen bonds between the drug and the O-2 atoms of thymine form the basis of its specificity for poly d(AT) however it is not clear from this work why TpA sequences should be preferred to ApT sequences.

Finally, it is clear that although TANDEM is very effective at preventing the de novo synthesis of poly d(AT) it is still necessary to check the products of a DNA polymerase catalyzed reaction for the presence of other contaminating DNA's. This was shown by the fact that E.coli DNA polymerase I will synthesize poly dA·poly dT if not occupied with the synthesis of another DNA.

Conclusion

De novo synthesis is a problem associated with replication of synthetic DNA's. DNA binding drugs of low specificity (such as netropsin) encourages the synthesis of DNA's of unusual sequence like poly d(TC)·poly d(GA). DNA binding drugs of high specificity (such as TANDEM) are useful inhibitors of poly d(AT) synthesis. The conformation that TANDEM recognizes in polyd(AT) also appears to exist in poly d(TAC)·poly d(GTA).

IV. Polypurine RNA Self-Structure

Introduction

The dismutation of polypyrimidine·polypurine DNA's under acid conditions results in the formation of a triplex and releases a fourth polypurine strand (Lee, Johnson and Morgan, 1979).

$$2[\text{poly d(TC)}\cdot\text{poly d(GA)}] \rightarrow \text{poly d(TC)}\cdot\text{poly d(GA)}\cdot\text{poly d}(C^{+}T) + \text{poly d(GA)}.$$

This polypurine strand was subsequently shown to form an ordered structure under neutral conditions which we have proposed to be four stranded (Figure 4-1)(Lee, Evans and Morgan, 1980). The evidence in favour of this hypothesis was:

1) Analogy with the fibre-diffraction deduced structures of poly rG and poly rI.

2) The failure of poly d(G m^6A) to form such complexes.

3) CD spectral features similar to those of poly dG.

Poly d(GAA), poly d(GA), poly d(GGA), poly d(IA) and poly d(G n^2A) were all shown to form such complexes with T_m's dependent on base composition. The first project I undertook was to investigate whether polypurine RNA's also formed analogous structures.

The Synthesis of Polypurine RNA's

Synthetic RNA's can be prepared by RNA polymerase transcription of the appropriate synthetic DNA's (Morgan, 1970). Figure 4-2 shows the kinetics of transcription. Typically there is a rapid burst of synthesis followed by a much slower reaction. In 24 hours in a number of different RNA preparations 0.9 ± 0.3 net fold synthesis (RNA phosphate: DNA

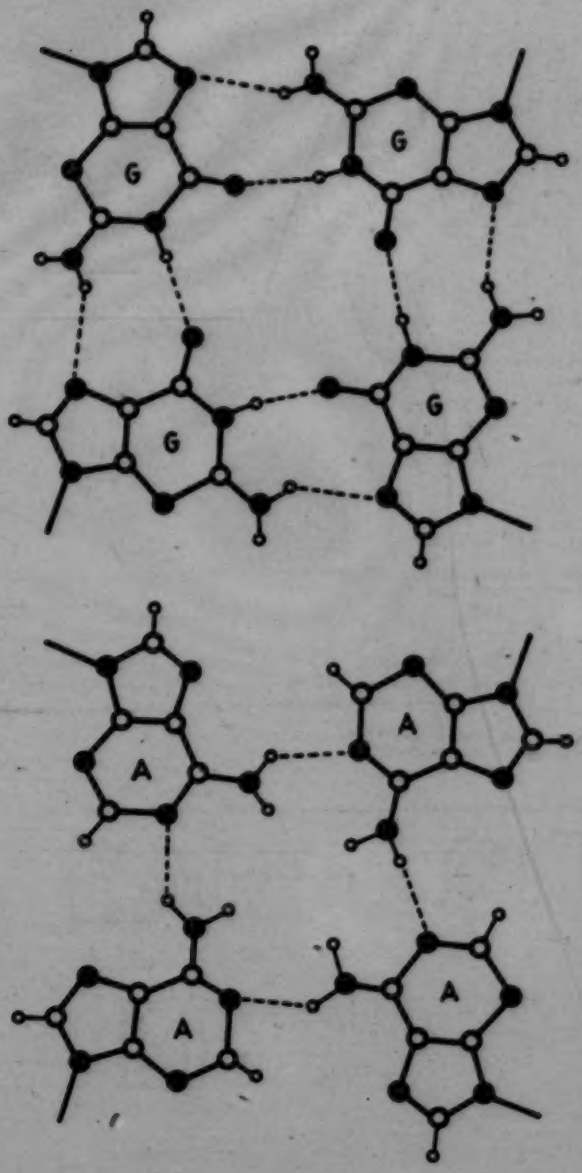


Figure 4-1 The two isomorphous base-tetrads.

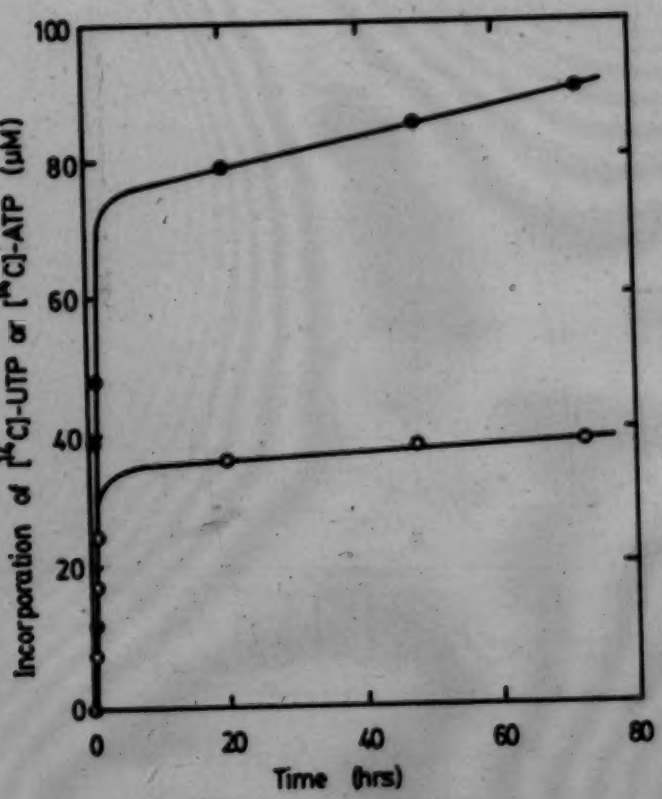


Figure 4-2 Kinetics of transcription. Conditions were as described in Chapter II with 170 μM poly d(TC)·poly d(GA) and 0.19 mg/ml RNA polymerase. Polypyrimidine synthesis (-o-): 1 mM UTP (20 μCi/ml [^{14}C]) and 1 mM CTP. Polypurine synthesis (-●-): 1 mM ATP (10 μCi/ml [^{14}C]) and 1 mM GTP. Samples (10 μl) were acid precipitated and counted as described earlier.

template phosphate) was observed. This poor net synthesis was a nuisance because of the high levels of DNA that had to be removed in subsequent purification steps. A variety of experimental modifications were examined in an attempt to improve yields without much success. The only factors consistently found to improve yields were the addition of either gene 32 DNA binding protein or single stranded DNA's complementary to the RNA being synthesized. Neither approach was of much practical use because improvements in yields were stoichiometrically dependent on these factors. The problem appeared to be intrinsic to the kinds of DNA's being transcribed. Many fold transcription of poly d(AT) would occur for example.

The most likely explanation for these observations is that described earlier (Morgan, 1970). The initial burst of synthesis is caused by RNA polymerase carrying out a first round of transcription subsequent synthesis is then dependent upon slow decomposition of DNA·RNA·RNA polymerase complexes. Figure 4-3 illustrates these points. Under conditions where the ratio of template molecules to polymerase molecules is greater than one net RNA synthesis is simply proportional to the amount of polymerase present. Under conditions where the ratio of template molecules to polymerase molecules is less than one net synthesis becomes essentially polymerase independent and the initial burst of synthesis appears to be circa 0.5 fold relative to the initial DNA concentration. Under both conditions a DNA·RNA·RNA polymerase complex is produced as illustrated on the following page. If this complex only very slowly decomposes and is not transcribed by other free polymerase molecules this explains the poor net synthesis.

My results suggest that these complexes do not participate in

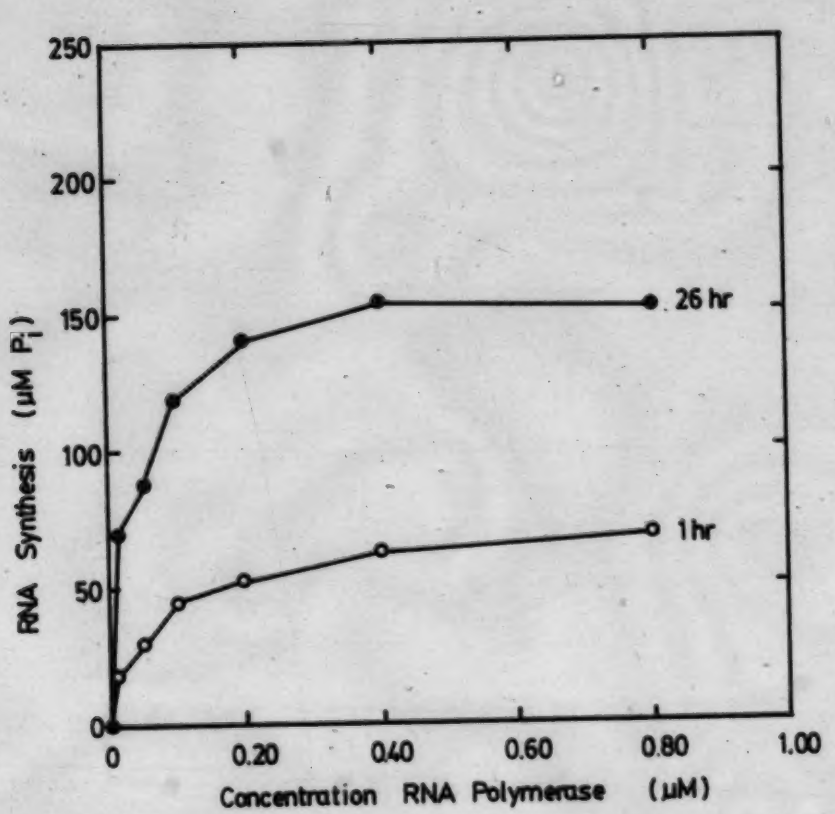
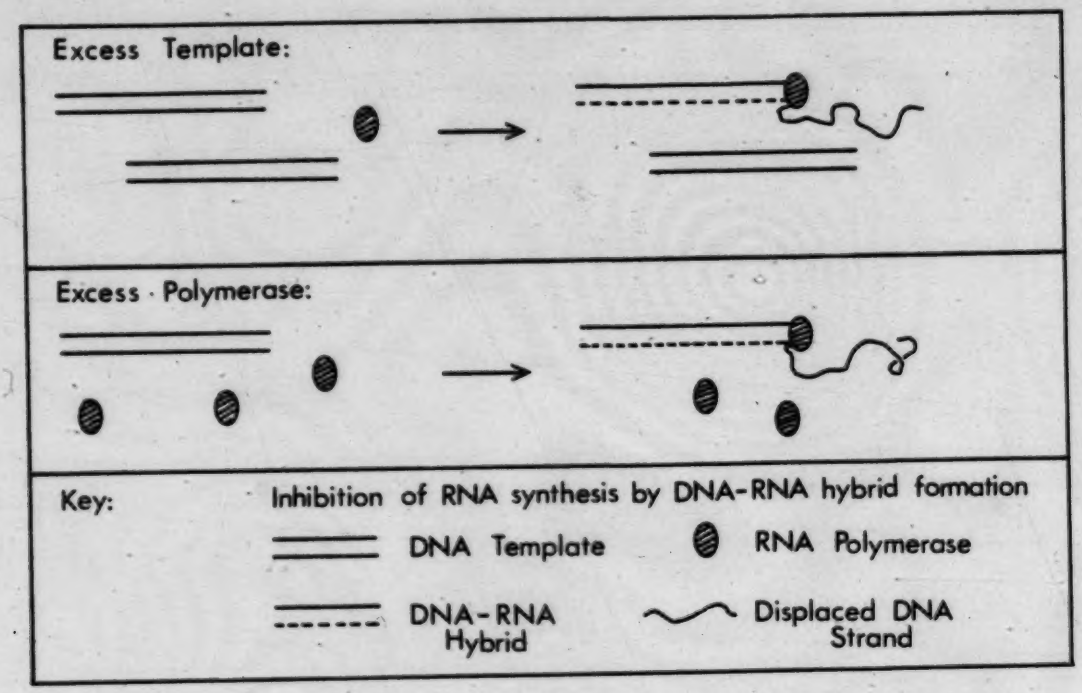


Figure 4-3 Effect of RNA polymerase to DNA strand ratio on net RNA synthesis. Conditions as described earlier with 1 mM ATP (0.56 μCi/ml) [^{3}H]), 1 mM GTP, 115 μM poly d(TC)·poly d(GA) (strand concentration 0.20 μM) and the indicated RNA polymerase concentrations. Samples (10 μl) were acid precipitated and counted. Net synthesis was actually quantitated by fluorescence measurements after annealing to excess poly d(TC).



further rounds of transcription. This is in contrast to earlier observations (Morgan, 1970) where in the presence of excess polymerase transcription continued at a significant rate after the burst phase. It still isn't clear why this should be. It probably reflects some difference between the RNA polymerases which were prepared by somewhat different methods. Since poor net synthesis didn't seem to be amenable to experimental manipulation further efforts were directed at purifying away the template DNA.

Purification of Synthetic RNA's

Digestion of DNA-RNA mixtures with DNase I was initially used to remove the DNA. Subsequently I found that while the total duplex fluorescence was destroyed rapidly, degradation of a radioactively labelled polydeoxypurine strand was a much slower process. RNA's prepared by a single DNase I digestion were found to contain essentially no complementary DNA but as much as 40% polydeoxypurine contaminant.

A number of methods were tried before a satisfactory procedure was discovered that would remove this contaminant. The resistance of the polypurine DNA strand is peculiar but is readily explained if one again assumes that the products of an RNA polymerase reaction are a DNA·RNA hybrid and displaced polydeoxypurine strand. The displaced strand will be DNase resistant since DNase I is specific for double stranded polynucleotides. DNase I has been reported to make the RNA strand of a DNA·RNA hybrid accessible to RNase (Wang, 1979). This suggests it can attack the DNA strand of a DNA·RNA hybrid which also accounts for the fluorescence assay data.

Exonuclease I is a single-strand specific 3' → 5' DNA exonuclease (Lehman and Nussbaum, 1964). Digestion of DNA·RNA mixtures specifically destroyed the DNA contaminant resulting in a product with approximately 2% residual DNA. Unfortunately a residual RNase activity was also apparent in this enzyme even after further purification and this made the technique less than ideal for routine preparation of RNA's of sufficient length to give good T_m's.

Cis - diols such as are found at RNA 3'-termini will form covalent adducts with dihydroxy boryl groups (Weith et al, 1970). The stability of such adducts is salt and pH dependent being stable at pH 8 to 9 at high (50 mM) $MgCl_2$ concentrations but unstable at lower pH's in the absence of magnesium (McCutchan et al, 1975). Although originally used for fractionating tRNA's this also provides the basis for separating DNA from RNA since DNA will not be retained on borate columns. Figure 4-4 illustrates such a column profile in which a DBAE-cellulose column was loaded with a sample originally containing 40% poly d(GA) and 60% poly r(GA). Following a flow-through peak a second peak was

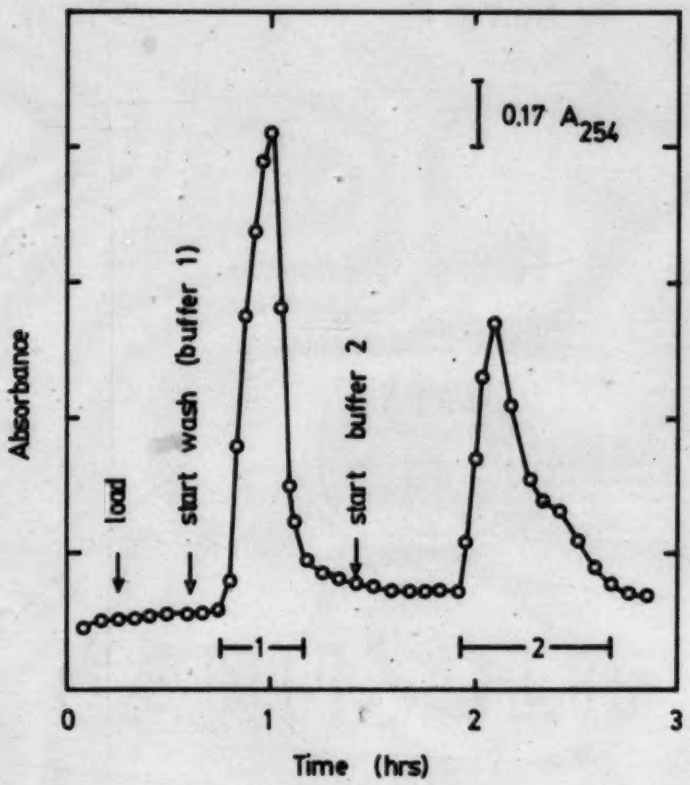


Figure 4-4 Purification of poly r(GA) by DBAE-cellulose chromatography. 3 ml poly r(GA), poly d(GA) mixture (40% d(GA), 60% r(GA)) was loaded in Buffer 1 by gravity onto a 5 ml DBAE-cellulose column then washed with Buffer 1 at 10 ml/hr. After eluting the flow-through the wash was changed to Buffer 2 at 10 ml/hr. Overall recovery of poly r(GA) (relative to amount synthesized) was 28% in this experiment.

observed to elute only upon lowering the pH and $MgCl_2$ concentration. T_m's showed no detectable DNA contamination. In a parallel experiment this purification was quantitated using radioactive DNA. The sample was initially contaminated to the extent of 8.7% poly d(GGA) in poly r(GGA). After chromatography the RNA in peak two contained only about 0.6% DNA. RNA in the column flow-throughs is probably due to the presence of some 2' or 3' phosphate terminated RNA's possibly the products of RNase activity*.

The DBAE-cellulose columns produced the cleanest products. Unfortunately irreproducible (and sometimes large) losses associated with the multiple chromatography steps also made this methodology less than ideal. Suspecting that the DNase I resistance of polypurine DNA's was due to formation of DNA·RNA hybrids I tried heat denaturing the DNase resistant material as shown earlier (Figure 2-3). This simple protocol dramatically improved the DNase sensitivity of polydeoxypurines. Using this method RNA's containing between 2 and 5% DNA were obtained quickly with only two chromatographic steps.

In general RNA's prepared by any of these (or a combination of these) methods were used in the following experiments.

Studies on Polypurine RNA's

Polypurine RNA's melt cooperatively with a T_m dependent on base composition (Figure 4-5(a), 4-5(b)) just as do polypurine DNA's. In buffers of high ionic strength these transitions are freely and rapidly reversible. The hyperchromicity varies somewhat depending on base composition (poly r(GAA) > poly r(GA) > poly r(GGA)) unlike polypurine

* Subsequent to developing this procedure a paper appeared (Ackerman et al, 1979) describing an identical application for DBAE-cellulose columns.

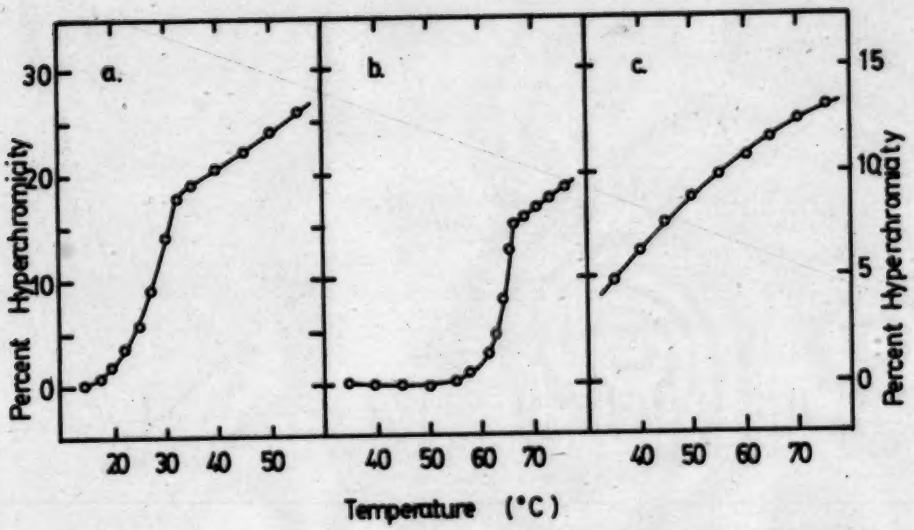


Figure 4-5 Thermal denaturation profiles. Panel (a) poly r(GAA); panel (b) poly r(GA); panel (c) poly r(G,A). In 250 mM NaCl, TE followed at 260 nm. Note different scales.

DNA's which normally show about 10% hyperchromicity independent of sequence. The reasons for this aren't obvious but do suggest some differences in stacking geometry between polypurine DNA's and RNA's. Polypurine RNA's of random sequence unstack during heating but cooperative melting isn't observed (Figure 4-5(c)). Table 4-1 summarizes the melting temperatures of these polypurine RNA's; polypurine DNA data is recorded for comparison.

Circular dichroism measurements confirm these observations. Figure 4-6 shows the CD spectra of poly r(GAA). The complex is characterized by strong ellipticity centred at 282 and 252 nm which disappears upon melting leaving weaker peaks at 270 and 250 nm. The ordered spectra are similar in shape to that of poly rG (Thiele and Guschelbauer, 1971) except that it is red-shifted by about 20 nm. The denatured spectra are almost identical to poly rA (Brahms et al, 1966) which is single stranded at this pH and salt. In general similar spectral changes were also associated with polypurine DNA melting.

Attempts to Form a poly rA Tetraplex

Considering that polypurine nucleic acids containing guanine residues form ordered structures it was interesting to see if poly rA would also form such structures at neutral pH's. Raising the ionic strength at pH 8 caused poly rA to form insoluble aggregates as could be followed by an increase in light scattering at 310 nm. Aggregation was also associated with a drop in 250 nm absorbance. These solutions cleared when heated but cooperative transitions weren't observed. Attempts to prepare soluble complexes by slow dialysis of salt into poly rA solutions or by nuclease treatment were unsuccessful. (It was hoped that mild treatment with a single-strand specific nuclease might break

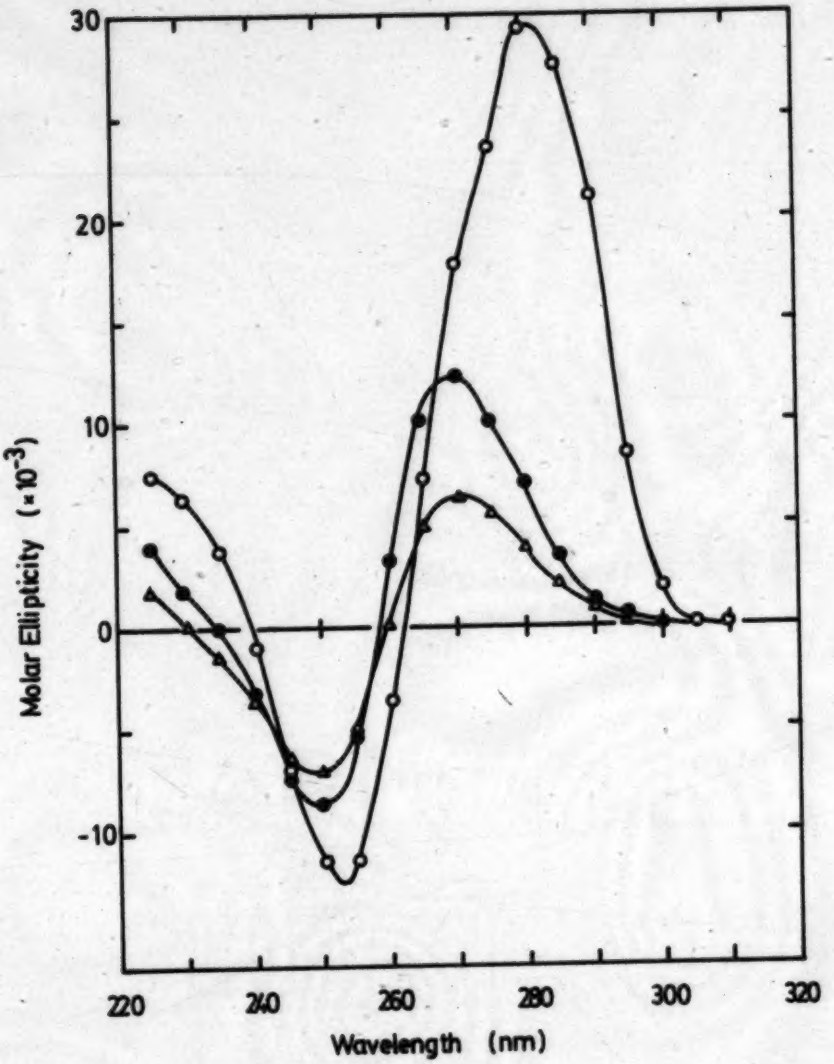


Figure 4-6 Poly r(GAA) CD spectra. (-o-) recorded in 250 mM NaCl, TE at 15°; (-●-) recorded in 250 mM NaCl, TE at 42°; (-Δ-) recorded in TE only at 16°. The molar extinction coefficient was taken to be 7500 (Murray, 1972). The spectrum after cooling back to 15° in 250 mM NaCl, TE was identical to that before heating.

Table 4-1

Melting temperatures of polypurine nucleic acids in 250 mM NaCl, TE.

RNA	T_m(C°)	DNA	T_m(C°)
poly r(GAA)	27	poly d(GAA)	45
poly r(GA)	64	poly d(GA)	61
poly r(GGA)	84	poly d(GGA)	69
poly r(G)	>100	poly d(G)	>100

Poly rG and poly dG data is from Englander et al (1972) and Miles and Frazier (1964) respectively.

up aggregates by nicking at branch points.) These experiments were discontinued because of these difficulties but do suggest some structure can be formed by poly rA in high salt. Somewhat surprisingly although many papers have been published concerning poly rA structure (eg. Holcomb and Tinoco, 1965; Stevens and Rosenfeld, 1966, etc.) only one has commented on this propensity to precipitate at neutral pH's in high salt (Leng and Felsenfeld, 1966). Poly dA has also been investigated (Riley, Maling and Chamberlin, 1966) even in high salt no cooperative melting behaviour is observed.

Discussion

These data suggest that polypurine RNA's form complexes analogous to those formed by polypurine DNA's. As indicated earlier most of this data is consistent with a four-stranded ("tetraplex") structure originally proposed for poly G and poly I. Subsequent to publishing these conclusions (Lee, Evans and Morgan, 1980) Gray et al (1980) published a reinterpretation of the poly d(GA) CD data concluding that while the structure was probably four-stranded the adenine residues were not incorporated into the helix. Rather the adenines were proposed to exist in a looped out configuration around a poly dG core. Such an interpretation is inconsistent with the failure of poly d(Gm^6A) to form a tetraplex and furthermore is inconsistent with one other simple experiment. As is shown later (Chapter V) it is possible for extrahelical residues to exist in perfectly stacked continuous duplexes.Thus there is no obvious reason why a mixed hybrid containing both, say, poly d(GA) and poly d(GGA) cannot form if the extrahelical residue hypothesis is correct. Figure 4-7 shows the effect of mixing poly d(GA) and poly d(GGA) together, denaturing, annealing and then melting. Each polymer

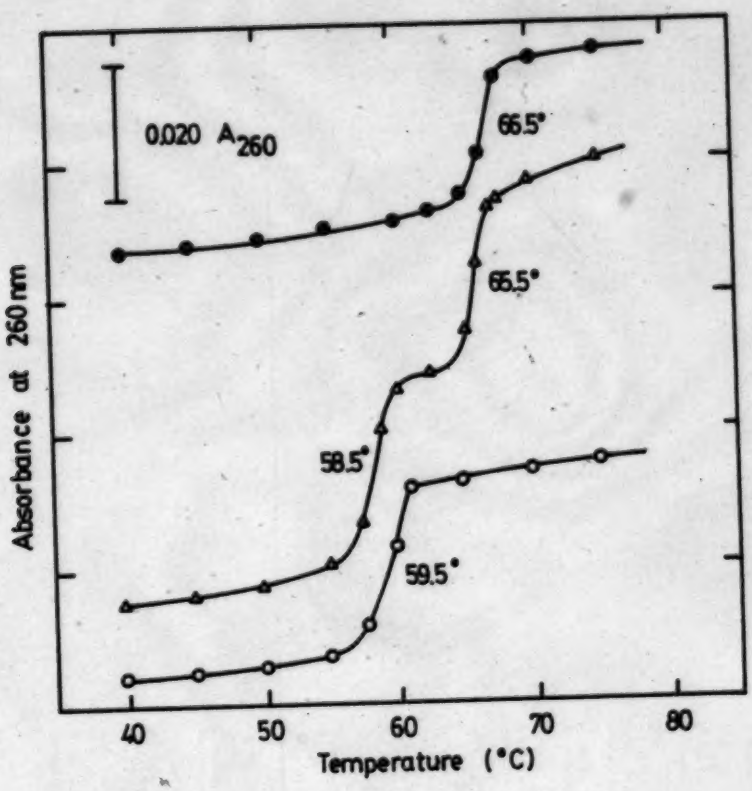


Figure 4-7 Poly d(GA) and poly d(GGA) will not form a hybrid of mixed composition. 300 µl samples contained either poly d(GA) (-o-), poly d(GGA) (-●-) or poly d(GA) and poly d(GGA) (-△-) 250 mM NaCl, TE. Samples were heat denatured (100°, 2 min) and cooled rapidly on ice. Samples were then melted and the absorbance followed at 260 nm. Initial absorbances were poly d(GA) 0.119 A_{260}, poly d(GGA) 0.129 A_{260}, and both together 0.218 A_{260}.

melts independently of the other without any trace of a mixed hybrid which we take to be evidence against the extrahelical hypothesis.

Conclusion

Polypurine DNA's form a structure that appears to be four stranded. Polypurine RNA's appear to form analogous structures.

V. Hybrid Polynucleotide Structures Containing Extrahelical Bases.

Introduction

The existence of extrahelical bases has long been postulated (Fresco and Alberts, 1960), and in many hypothetical secondary structures of various RNA's extrahelical bases are often drawn to maximize Watson-Crick base-pairs (Tinoco et al, 1973; Salser, 1977).

Lomant and Fresco have extensively investigated the properties of homopolymer RNA duplexes in which one strand contains a variable amount of a non-complementary base. As a result of mixing curve analysis (Lomant and Fresco, 1972) and photochemical reactivity (Lomant and Fresco, 1973; 1975) they concluded that hybridization of complementary polyribonucleotides with non-complementary bases inserted into one strand usually results in these bases taking up an extrahelical conformation while the complementary bases form an otherwise perfect Watson-Crick duplex. Bases capable of forming "wobble" pairs were found to represent a significant exception to this generalization (Lomant and Fresco, 1972).

More recently high resolution NMR has provided further evidence for the existence of extrahelical bases. In small single-stranded oligonucleotides strong stacking interactions can lead to a "bulging-out" of non-stacking bases. In $m_2^6AUm_2^6AU$ for example the A's are stacked with the intervening U in a conformation analogous to an extrahelical base (Hartel et al, 1981). In another study on the effects of internal non-Watson-Crick base oppositions on the stability of a large self-complementary oligodeoxyribonucleotide ATCCTA[T_n]TAGGAT, Haasnoot et al (1980) demonstrated that short internal loops (n = 1,2) could be stabilized by the surrounding sequence.

Table 5-1 summarizes some of the hybrid polymers in which extrahelical bases have been observed. A variety of modified bases, incapable of base-pairing, have also been observed to take up an extrahelical configuration. It should be noted that a significant number of these randomly inserted non-complementary bases will appear in runs of two or more. None of these studies have been repeated with DNA or defined repeating sequence polymers.

Such structures are not simply theoretical or biophysical curiosities. As described earlier they are found in tRNA's. In procaryote 5s rRNA's an extrahelical adenine at position 66 occurs universally and is distinguished by its sensitivity to nuclease and chemical modification (Garrett et al, 1981 ; Hori and Osawa, 1979). It has been postulated to be part of the L18 binding site (Garrett et al, 1981) and analogous structures are observed in eucaryote 5s rRNA's. Finally it has been shown that spontaneous and mutagen induced frame-shifting mutations often originate in "hotspots". These are typically located in base runs (eg. -AAAAA-) and it is easily seen how insertions and deletions can be generated during repair/replication by transient extrahelical bases (Drake and Baltz, 1976; Streisinger et al, 1966).

While studying the properties of synthetic repeating sequence nucleic acids I found that many apparently non-complementary single stranded nucleic acids can form stable complexes at neutral pH's and moderate ionic strengths. In this Chapter a number of these complexes are characterized as being duplexes with single extrahelical bases and some of their physical properties are considered.

Hybrids Composed of poly d(GGA) and poly d(TC)

Mixing curve analysis has proved to be one of the most

Table 5-1

RNA hybrids containing extrahelical bases [from Lomant and Fresco, (1975)].

Polymers	extrahelical base
poly (A,U) + poly U	U
poly (G,U) + poly C	U
poly (I,C) + poly C	C
poly (I,U) + poly C	U
poly (U,A) + poly A	A
poly (C,I) + poly I	I

useful techniques available for studying extrahelical bases (Lomant and Fresco, 1972). The principle behind the technique is that extrahelical bases will often be indicated by a distinctive strand stoichiometry. In the case of duplexes containing all purines in one strand and all pyrimidines in the other this will also be reflected by a distinctive pyrimidine to purine base ratio. Mixing curves were obtained in two different ways. The first (Figure 5-1(a)) uses the method of continuous variations. The stoichiometry of the complex was 43 ± 2% mole fraction purine (1.33 ± 0.10 pyrimidine per purine). The second technique takes advantage of the observation that ethidium bromide fluoresces in the presence of hybridized strands but shows no fluorescence in the presence of truly single strands. By adding a variable amount of poly d(TC) to a fixed amount of poly d(GGA) a break was seen at a base ratio of 1.34 ± 0.03 pyrimidines to 1.00 purines (Figure 5-1(b)) a result in excellent agreement with the first method.

Neither result was significantly different from 1.33 pyrimidine to 1.00 purine, this immediately suggested a structure of the type diagrammed below.

```
-GGAGGA-
 ......
-CCTCCT-
  \/  \/
   T   T
```

Such a structure contains 4 pyrimidine bases for each 3 purines, the extra thymidine being forced out into an extrahelical configuration. Using CPK models a well-stacked duplex could be built containing extrahelical pyrimidines. The structure appeared to be sterically reasonable and stacking was continuous even adjacent to the extrahelical bases (Figure 5-2).

Examination of this stacking suggested the experiment illust-

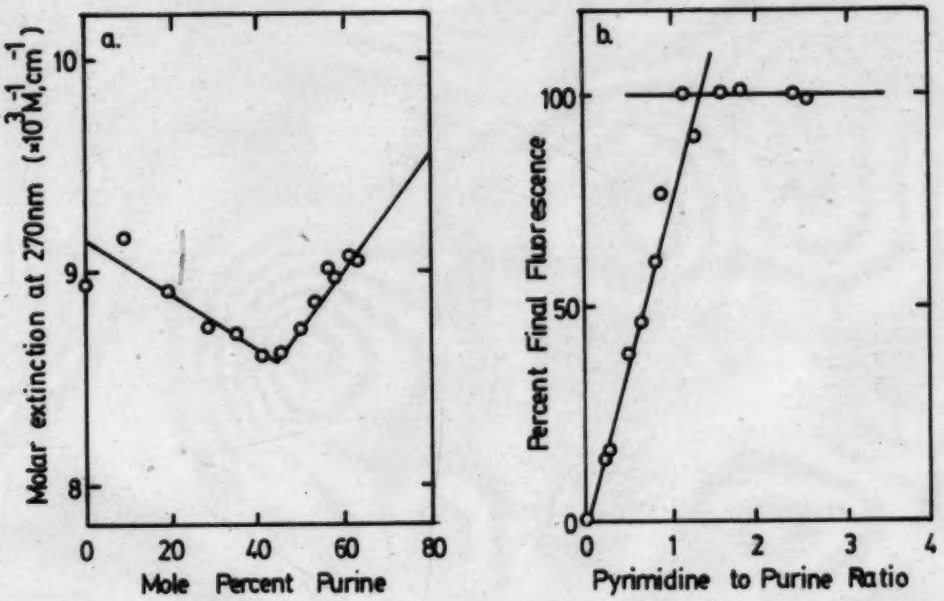


Figure 5-1 Determination of strand stoichiometry by mixing curve analysis. (a) By UV absorbance. Poly d(GGA) (697 µM) was added in successive small aliquots to 300 µl of 52.0 µM poly d(TC) in 0.5 M NaCl, 0.1 M Tris·HCl pH 8.0, 0.1 mM EDTA. After annealing (1.0 hours at 22°) A_{270} was determined, corrected for dilution and ε_{270} calculated based upon known polynucleotide concentrations. A base composition of 43 mole percent purine is indicated. (b) Hybridization measured by ethidium bromide fluorescence. Aliquiots (40 µl) containing 0.93 nmol poly d(GGA), 0.5 M NaCl, 0.1 M Tris·HCl pH 8.0, 0.1 mM EDTA and 0-2.4nmol poly d(TC) were annealed 1.0 hour at 22°. 2.0 ml pH 8.0 ethidium bromide solution was then added and fluorscence measured as described earlier. A 1.34:1 pyrimidine to purine ratio is indicated.

Figure 5-2 A CPK model of the poly d(GA)·poly d(TTC) hybrid. Note that the bases adjacent to extrahelical residues are still well stacked.

rated in Figure 5-3. Since the cytosines bracketing the extrahelical thymidine appeared to be well-stacked they should be photodimerizable with ultraviolet light. Figure 5-3 shows this prediction to be correct. In this experiment DNA containing radioactive pyrimidines was UV-irradiated and the photoproducts were analysed by formic acid hydrolysis of DNA followed by paper chromatography (Setlow and Carrier, 1966). Controls showed that no $\hat{C}\hat{C}$ dimers were formed when poly d(TC) annealed to the correct strand was irradiated with 254 nm light (Figure 5-3(a)) nor were any formed when single-stranded poly d(TC) was irridiated (Figure 5-3(b)). However under conditions such that a hybrid between poly d(TC) and poly d(GGA) formed, irradiation produced a peak of radioactivity corresponding to a $\hat{C}\hat{C}$ derived $\hat{U}\hat{U}$ dimer (Figure 5-3(c)). This provides strong evidence in favour of the proposed structure and points to the minimal distortion found around each extrahelical base.

Circular dichroism was also used to investigate other features of the hybrid structure. Upon formation of this complex the CD spectrum showed a pronounced decrease in ellipticities at 240 and 276 nm and the loss of a shoulder at 260 nm (Figure 5-4). The observed spectrum could be modelled quite well by adding together the CD spectra of poly d(GGA)·poly d(TCC) (the "duplex" portion of the hybrid) and of 5'-dTMP (the "extrahelical" portion) in appropriate ratios. The agreement between model and observed spectra again indicates how little the presence of extrahelical bases perturbs the core duplex. Ultraviolet absorbtion spectra were also recorded. Upon complex formation hypochromicity was observed (Figure 5-5). The magnitude of this hypochromicity was not as great as that calculated by adding together the absorbtion spectra of poly d(TCC)·poly d(GGA) and 5'-dTMP indicating, contrary to CD measure-

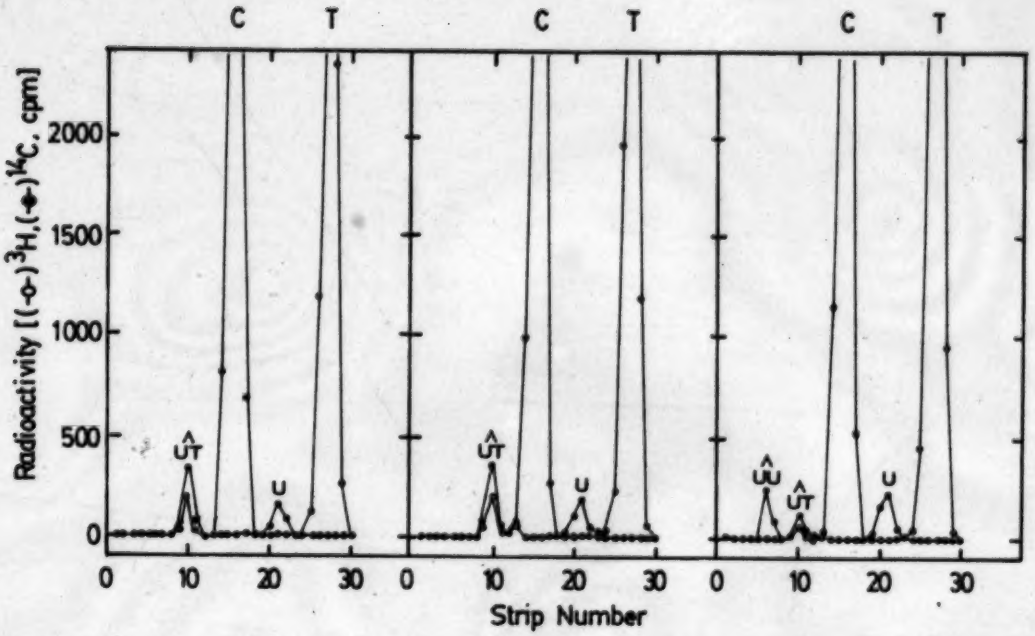


Figure 5-3 Evidence for the exist nce of extrahelical bases as shown by the production of unexpected pyrimidine dimers. Samples (400 µl) contained 7.1 nmol doubly-labelled poly d(TC)(1500 cpm·nmol^{-1} ^{3}H-thymidine, 2300 cpm·nmol^{-1} ^{14}C-cytidine), a two-fold excess of the indicated purine strands, 0.5 M NaCl, 0.1 M Tris·HCl pH 8.0 and 0.1 mM EDTA. Samples were irradiated with 6.0×10^4 erg·mm^{-2} 254 nm UV light in 1 mm pathlength quartz cuvettes, ethanol precipitated in the presence of carrier DNA, hydrolysed and the bases chromatographed (17 hours) on Whatman Number 1 paper. Panel (a) is a control containing poly d(TC) annealed to poly d(GA). Only one pyrimidine dimer was produced with an R_f and labelling characteristic of ÛT (ĈĈ and ĈT dimers are deaminated to ÛU and ÛT respectively during the hydrolysis step). Panel (b) is a second control containing poly d(TC) and poly d(GGA), annealing was prevented by omitting NaCl. Again, only ÛT dimers are formed. Panel (c) shows poly d(TC) annealed to poly d(GGA). A second pyrimidine dimer appears with a labelling and R_f characteristic of ÛU.

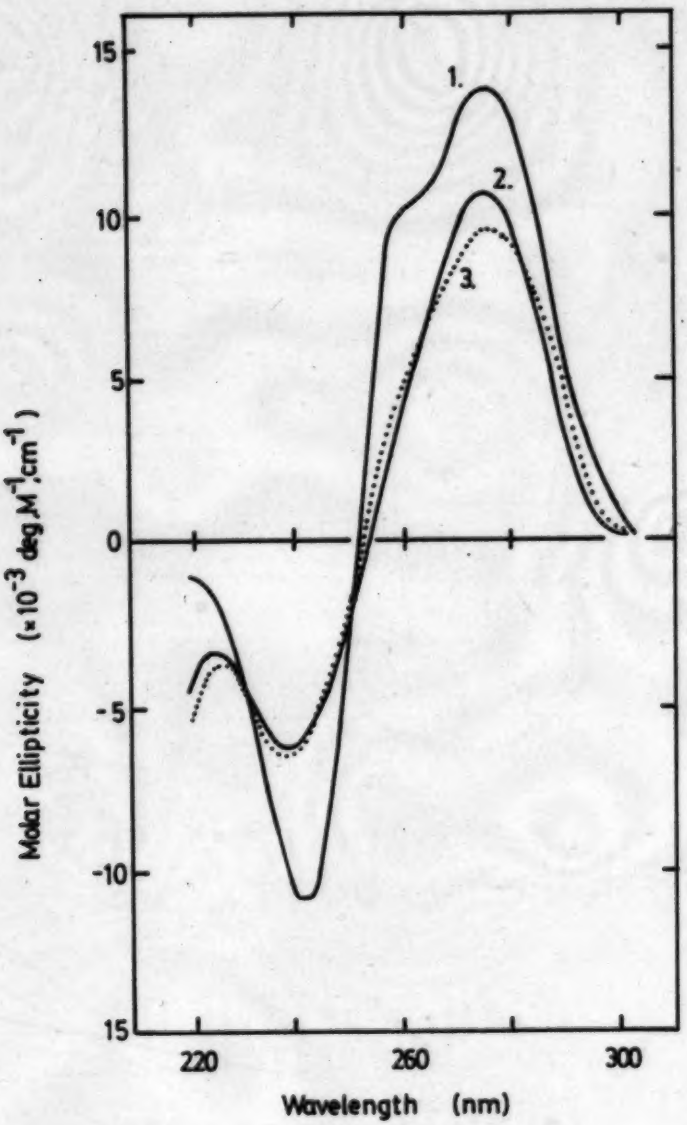


Figure 5-4 Spectroscopic evidence for the existence of extrahelical bases. Circular dichroism (CD) was measured in 0.5 M NaCl, 0.1 M Tris·HCl pH 8.0, 0.1 mM EDTA at 22°. Curve (1) is a hypothetical spectrum produced by adding the observed spectrum of isolated poly d(TC) to that of poly d(GGA) using the formula $\theta_{calc} = [1.33\ \theta_{pyr} + 1.0\ \theta_{pur}] \div 2.33$. Curve (2) is the observed spectrum of this hybrid. Curve (3) was modelled using the CD spectra of poly d(TCC)·poly d(GGA) and 5' dTMP and the formula $\theta_{calc} = [6.0 \times \theta_{duplex} + 1.0 \times \theta_{dTMP}] \div 7$.

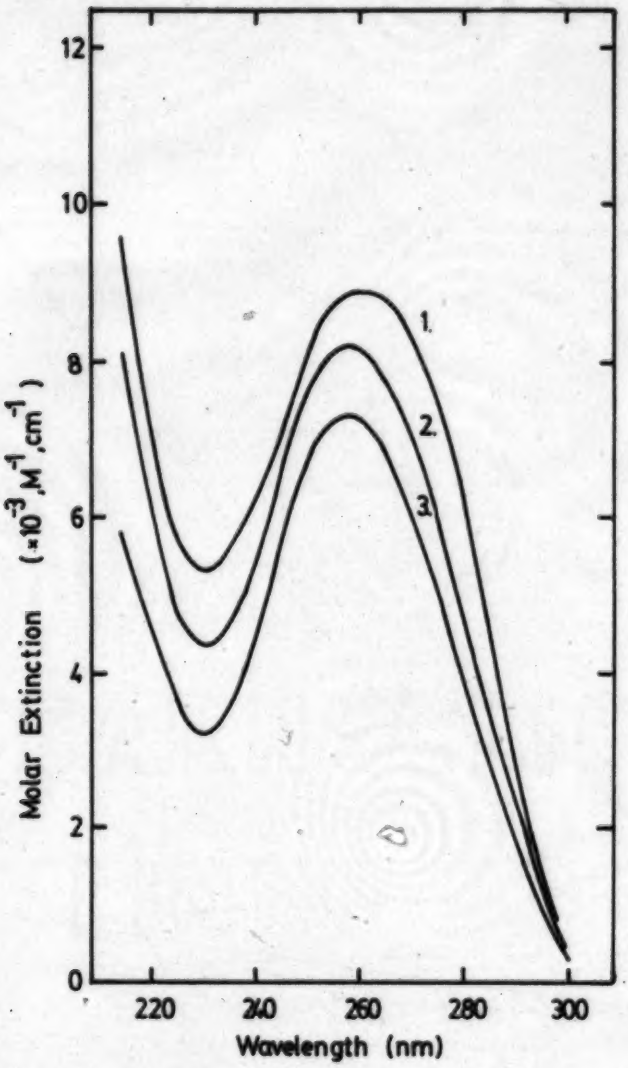


Figure 5-5 Demonstration of hypochromicity during complex formation. Conditions and labelling as in Figure 5-4. Note that the observed hypochromicity lies between the "non-interacting" and modelled curves.

ments, that the extrahelical bases do disrupt stacking to some extent. This is presumably a reflection of the fact that hypochromism is a more sensitive function of chromaphore separation (r^{-3}) than is CD (r^{-2}) (Cantor and Schimmel, 1980). CD is normally very sensitive to changes in chromaphore orientation. The observation that the hybrid spectra can be modelled using duplex and 5'-dTMP spectra suggests (but does not prove) that the relative rotation of base-pairs surrounding an extrahelical base closely resembles the relative rotation in the duplex.

Hybrid melting could be followed spectrophotometrically at 260 nm. the hybrid melts cooperatively and is reasonably stable. A T_m of 47.5° (Table 5-2) was found in 0.10 M NaCl, 10 mM Tris HCl pH 8.0 0.1 mM EDTA, which is 40° lower than that of poly d(TCC) poly d(GGA) (89°) (Table 5-3) in the same buffer.

The hybrid T_m is much more sensitive to cation concentration than are ordinary pyrimidine-purine DNA's (19.1° per ten-fold change in NaCl concentration verses 13±3). Since electrostatic repulsion is a function of charge separation this increased salt dependence suggests that on average phosphate residues are closer to one another in the hybrid than in ordinary duplex DNA's an observation in qualitative agreement with CPK model building studies.

Other Hybrids Containing Extrahelical Bases

Considering the information gained from studying the poly d(GGA) poly d(TCTC) hybrid a number of other polymer combinations were examined to see which, if any, would hybridize and if so which also contained extrahelical bases.

Table 5-2

Melting points of hybrids containing extrahelical bases in 0.10 M NaCl, TE, determined by UV spectroscopy.

Hybrid	T_m (C°)	dT_m /d log [NaCl]	D(C°)
poly d(CGA)·poly d(TCTC)	47.5°	19.1	12.4
poly d(GA)·poly d(TTC)	29°	20.8	10.6
poly r(GAGA)·poly d(TCC)	54°	17	—
poly r(GAGA)·poly d(TTT)	32°	10	—

Table 5-3

Melting points of ordinary synthetic polynucleotides in 0.10 M NaCl, TE, determined by UV spectroscopy.

Polymer	T_m (C°)	$dT_m/d \log[NaCl]$
poly dA·poly dT	69.5	15.8
poly d(TTC)·poly d(GAA)	75.5	14.0
poly d(TC)·poly d(GA)	82	13.7
poly d(TCC)·poly d(GGA)	89	13.0
poly dG·poly dC	99.5	10.4
poly r(GAA)·poly d(TTC)	82	16
poly r(GA)·poly d(TC)	92	16
poly r(GGA)·poly d(TCC)	>100	--
poly d(TG)·poly d(CA)	87	--

Note: DNA's were prepared with DNA polymerase, DNA·RNA hybrids by transcription of the single stranded DNA's with RNA polymerase.

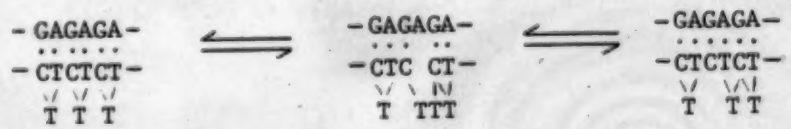
poly d(GA)·poly d(TTC)

Mixing curve analysis gave a strand ratio of 1.5 pyrimidines to 1.0 purine consistent with a structure containing every other thymidine in an extrahelical configuration (see below). CD spectra supported this hypothesis (Figure 5-6). The T_m of this hybrid was considerably lower than poly d(GGA)·poly d(TCTC) (Table 5-2) presumably because of both the higher proportion of extrahelical bases and the intrinsically lower T_m of the core duplex [poly d(TC)·poly d(GA) vs poly d(TCC)·poly d(GGA)].

Pyrimidine dimers were also examined (Figure 5-7) since one would predict a decrease in $\hat{T}T$ dimer production in the hybrid. The observed order of $\hat{T}T$ dimer production was:

(most) poly d(TTC) > poly d(GA)·poly d(TTC) > poly d(GAA)·poly d(TTC)(least)

This is not too surprising if considering that this technique can photochemically trap intermediates one of which would be expected to contain transiently stacked thymidines ie.



Photochemical events such as pyrimidine dimer production take place extremely rapidly (10^{-12} sec; Setlow and Pollard, 1962) far more rapidly than conformational changes (10^{-4} - 10^{-6} sec). These observations suggest that these are dynamic structures with transiently unpaired intermediates analogous to those observed by NMR and hydrogen exchange. (Early et al, 1981a,b; Hogan and Jardetzky, 1980; Mandal et al, 1979)

There was no significant variation in the amount of $\hat{C}T$ dimer produced. This points out a disadvantage of dimer experiments. Al-

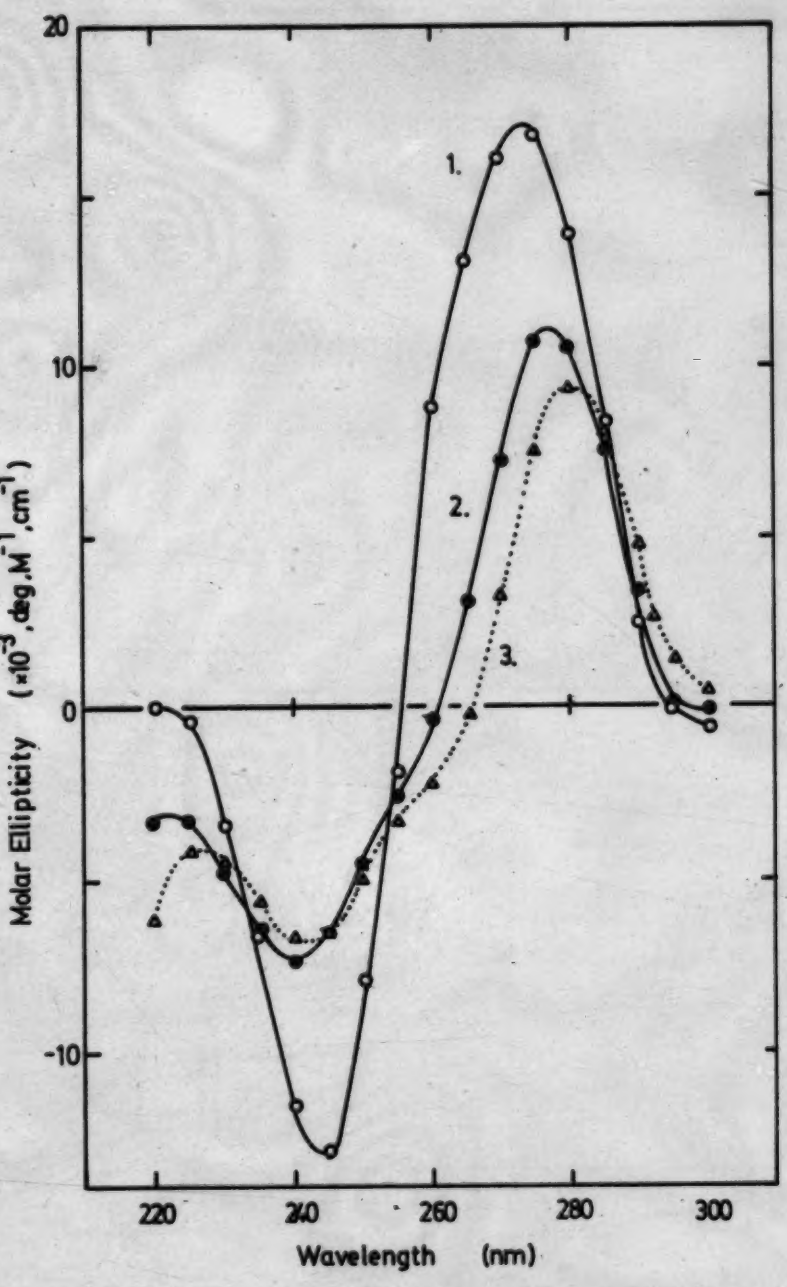


Figure 5-6 CD evidence for the formation of a hybrid between poly d(GA) and poly d(TTC). Conditions were as described in Figure 5-4. Curve (1) is a hypothetical spectrum calculated by adding the observed spectrum of poly d(TC) to that of poly d(GA) using the formula $\theta_{calc} = [1.5\ \theta_{pyr} + 1.0\ \theta_{pur}] \div 2.5$. Curve (2) is the observed spectrum of this hybrid. Curve (3) was modelled using the spectrum of poly d(TC)·poly d(GA) and 5' dTMP and the formula $\theta_{calc} = [4\ \theta_{TC\cdot GA} + 1.0\ \theta_{dTMP}] \div 5$.

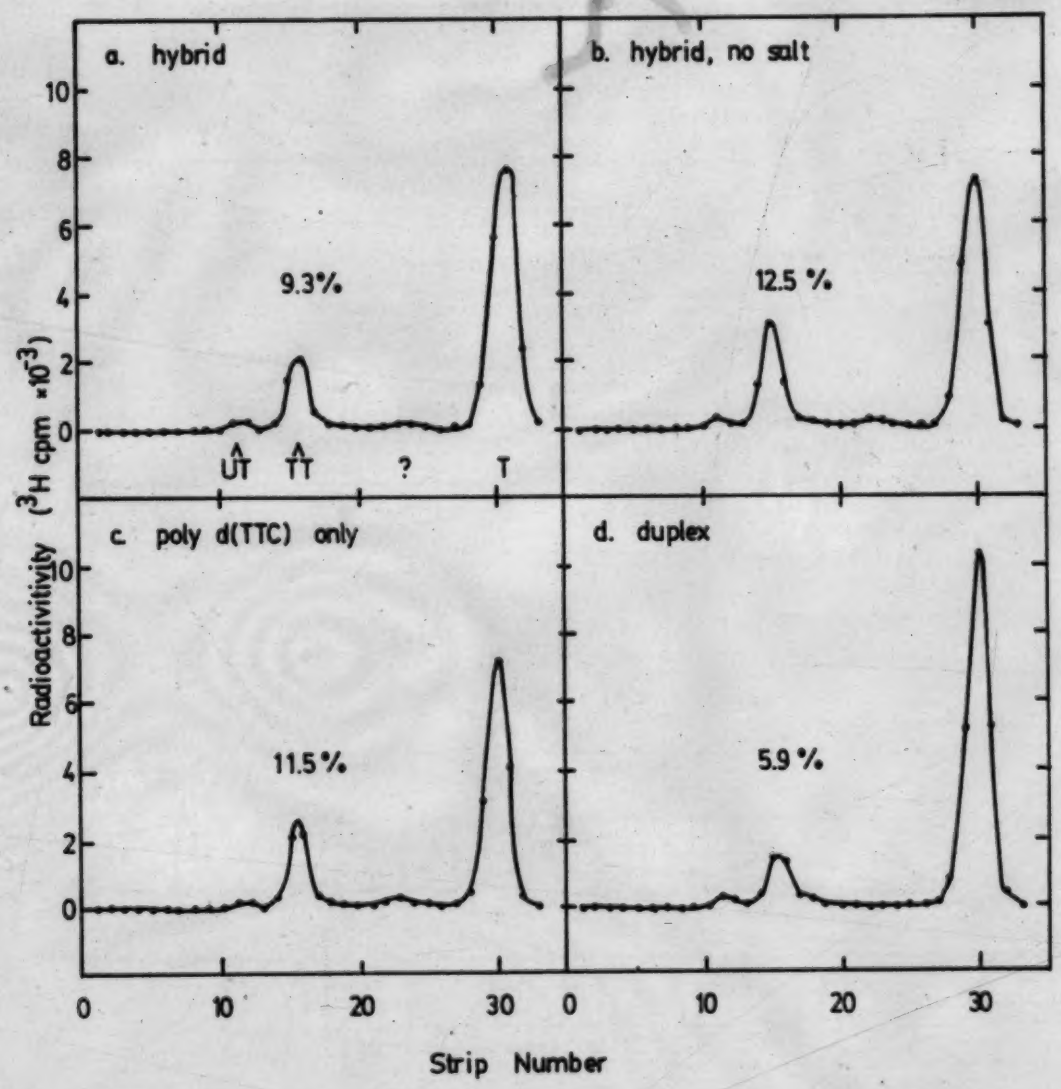


Figure 5-7 Production of pyrimidine dimers by irradiation of poly d(TTC). Samples contained singly labelled poly d(TTC) (3800 cpm·nmol^{-1} ^{3}H-thymidine) 8.7 nmol in 150 µl and a two-fold excess of the purine strand where indicated. Other conditions as in Figure 5-3. Panel (a) is the hybrid poly d(TTC) annealed to poly d(GA), 9.3 mole % of the thymines plus dimers appeared as T̂T dimers. Panel (b) is a control in which the salt was omitted to prevent annealing of the poly d(TTC) and poly d(GA), 12.5% of the thymines were found as dimers. Panel (c) contained poly d(TTC) only, 11.5% as dimers. Panel (d) poly d(TTC) annealed to poly d(GAA), 5.9% as dimers. No significant variation in ĈT production was observed (2.1 ± 0.3%).

though observation of dimer production is good evidence of a particular stacking arrangement, a negative experiment doesn't prove or disprove anything. Formation of cyclobutane rings is probably very sensitive to stacking geometry and one could easily imagine situations where the exact hybrid geometry disfavours dimerization. In these experiments I also occasionally observed small amounts of unidentified photo-products running at R_f's of 0.47 ± 0.02 (eg. Figure 5-7) and 0.53 [extensively irradiated poly d(TC)]. These ran close to uracil ($R_f = 0.50$) but the radioactivity suggested a thymine origin. They didn't seem to correspond to any of the known minor thymine (Varghese, 1970) or thymine-cytosine (Varghese, 1971) photo-products produced in frozen aqueous nucleoside solutions. Whether this has any relevance to the recent observation of Lippke et al (1981) is unknown.

poly r(GAGA)·poly d(TCC)

Hybridization of the ribopolymer poly r(GA) with poly d(TCC) gave a strand ratio (by fluorescence) of 0.75 pyrimidines to 1.0 purine consistent with a structure containing extrahelical adenines and a poly r(GGA)·poly d(TCC) core.

```
- TCCTCC -
  ......
- AGGAGG -
    V V
    A A
```

The T_m was slightly higher than poly d(GGA)·poly d(TCTC) which contains the same core sequence. This difference can be rationalized completely by observing that poly r(GGA)·poly d(TCC) is itself intrinsically more stable than poly d(GGA)·poly d(TCC). This mixing ratio was distinctly different from that found when poly d(GA) was annealled to poly d(TCC). (1:1). The possible structure of this hybrid is discussed in Chapter VI.

poly r(GAGA)·poly d(TTT)

Hybridization of poly r(GA) with poly dT also gave a 0.75 pyrimidine to 1.0 purine ratio. This was curious and the most probable structure is a hybrid consisting of both G·T pairs and extrahelical G's

```
- TTTTTT -
  ......
- GAAGAA -
    \/ \/
    G  G
```

an observation consistent with the relatively low stability of this structure (Table 5-2). In order to confirm this hypothesis advantage was taken of the fact that glyoxal (1,2-ethanedial) will reversibly form an adduct with guanine base. At pH ~7 this reaction is specific for guanine alone (Broude and Budowsky, 1971). The reaction can be followed at 268 nm spectrophotometrically and because the adduct involves the N-2 amino group and N-1 positions this provides a convenient way in which to titrate accessible and base-paired guanines. Figure 5-8 illustrates the kinetics of this reaction. Despite the large error bars (these experiments consumed rather large quantities of scarce poly r(GA) so absorbance had to be low) it seems that when hybridized only about half the guanines are accessible to glyoxal compared to the number accessible if the RNA is first hydrolysed to mononucleotides. Plotted as a first order function it is apparent that those guanines in the hybrid are also more reactive ($t_{1/2}$(hybrid) = 9.8 min, $t_{1/2}$(mononucleotides) = 25 min). This is easily explained considering that GMP stacks in such a way as to restrict access of reagent whereas extrahelical guanines are far more exposed. The inset illustrates that the hybrid is stable in the presence of 50 mM glyoxal with a T_m about what one would expect in this salt. After melting, however, the hybrid doesn't reanneal. Somewhat surprisingly poly d(GA) didn't anneal to poly dT.

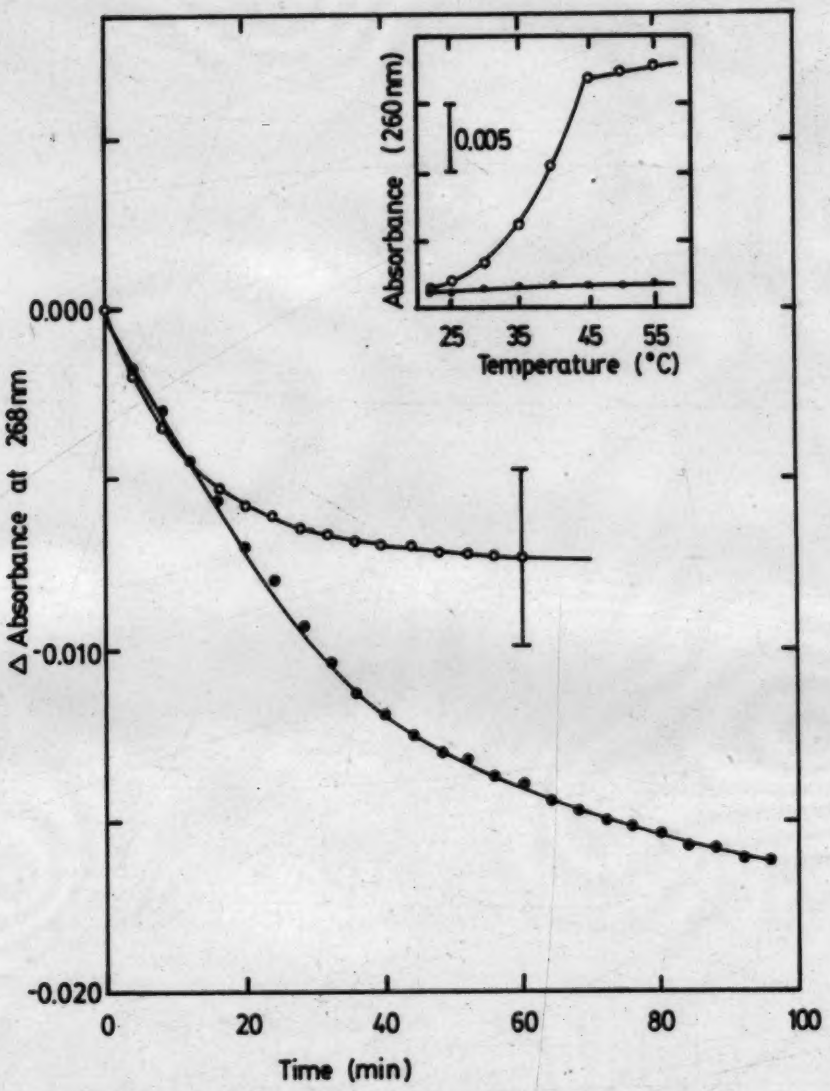


Figure 5-8 Reaction of poly r(GAGA)·poly d(TTT) with glyoxal. Final concentrations were (in 1.1 ml) 45 mM KPi, 0.45 M NaCl, 4.4 µM poly r(GA) (or hydrolysed poly r(GA)), 4.6 µM poly dT and 46 mM glyoxal, measured pH 7.0. Reaction progress was recorded automatically at 268 nm, room temperature in sectored quartz cuvettes (0.875 cm path length) and were initiated by mixing the separate solutions of polynucleotide and glyoxal. Poly r(GA) was hydrolysed by digesting with 0.10 M NaOH at 100° for 5' and neutralized with HCl. A dummy digest of the polymer was performed omitting the heating step. Control (-•-), hydrolysed poly r(GA) plus poly dT, 2 replicates. Hybrid (-o-), poly r(GA) annealed to poly dT, 3 replicates. Inset: Melting of glyoxal reacted poly r(GAGA)·poly d(TTT) hybrid. Poly r(GAGA)·poly d(TTT) after reaction was transferred to 1 cm quartz cells and melted. A T_m of 37° was observed (-o-). After storing on ice overnight the sample was remelted. No transition was observed (-•-).

Table 5-4

Probable structures of hybrid polymers containing extrahelical bases.

Hybrid	Measured pyr:pur ratio	Probable ratio	Structure
poly d(GGA)·poly d(TCTC)	1.34 ± 0.03:1	1.33:1	-GGAGGA- ..... -CCTCCT- \/ \/ T T
poly d(GA)·poly d(TTC)	1.48:1	1.1:1	-GAGAGA- ..... -CTCTCT- \/ \/ \/ T T T
poly r(GAGA)·poly d(TCC)	0.78:1	0.75:1	-TCCTCC- ..... -AGGAGG- \/ \/ A A
poly r(GAGA)·poly d(TTT)	0.78:1	0.75:1	-TTTTTT- ..... -GAAGAA- \/ \/ G G

Note: Strand ratios were determined as indicated in the legend to Figure 5-1(b).

Discussion

Are mixing ratios determined in the presence of ethidium bromide indicative of ratios in the absence of ethidium? It isclear in all these cases hybrids will form in the absence of ethidium since cooperative melting transitions were observed but are these the same structures? Intercalative drugs are known to stabilize some odd base pairs (eg. Westhof et al, 1980; Helfgott and Kallenbach, 1979) to alter the conformation of polynucleotides (eg. Patel and Canuel, 1979), and shift equilibria between polynucleotide structure (eg. Waring, 1974). They have also been hypothesized to stabilize extrahelical bases by stacking on the extrahelical base or by intercalation (Streisinger et al, 1966; Lee and Tinoco, 1978).

Nevertheless in the case of poly d(GGA)·poly d(TCTC) the strand ratios found in the presence of ethidium could be duplicated by UV-mixing experiments showing that ethidium need not be present to stabilize such structures. Furthermore no evidence in conflict with mixing ratios was observed. T_m's fit a rational progression and CD, UV and chemical modification experiments were consistent with extrahelical models. Thus although ethidium does stabilize these hybrids it does not appear to perturb the mixing ratios and its absence can be compensated for by a sufficiently high counterion concentration.

What are the major determinate errors present? These can be summarized as errors associated with the presence of contaminating DNA's, errors in extinction coefficients and errors associated with end effects. The first two types of error are minimized by the manner in which mixing stoichiometries are determined. For example by titrating polypurine preparations with polypyrimidine standards one measures directly the

free polypurine concentration and can ignore the residual levels of contaminating duplex. Furthermore since similar errors are likely present in all polypyrimidine extinction coefficients and these are always propagated into the polypurine concentration measurements many errors will simply cancel out when strand ratios are taken. This is presumably the reason why strand ratios came out well despite the fact that extinction coefficients are probably accurate to no more than 5 to 10%. End effects are a much more important source of error since T_m's are dependent upon strand length (Cassani and Bollum, 1969). Ordinary polymers of $n \simeq 50$ may have a T_m as much as 10° lower than polymers of infinite length (Martin et al, 1971; Tazawa et al, 1972) suggesting that the T_m's reported here should be considered in a relative rather than absolute manner. For polymers of $n \geq 50$ there is little dependence of T_m on concentration (Martin et al, 1971) and although mixing ratios have been reported to be sensitive to polymer length [Cassani and Bollum (1969) demonstrated that when $n < 12$ mixtures of dA_n and poly dT form triplexes in preference to duplexes and vice versa when $n > 12$] such effects are not likely to occur with polymers of the size used in these studies.

Is there any competition from the polypurine tetraplex? In Figure 5-9 two melting transitions can be seen. The first is the hybrid and the second is the tetraplex. When mixed at temperatures below the hybrid T_m mixing curves were fairly sharp showing that all the available polypurine strands were complexing with polypyrimidine strands. It is likely that the polypurine structure forms only after the hybrid melts and points out the danger of equating T_m and stability. Although the hybrid has a lower T_m it is the preferred state of the polypurine strand in the presence of a polypyrimidine strand.

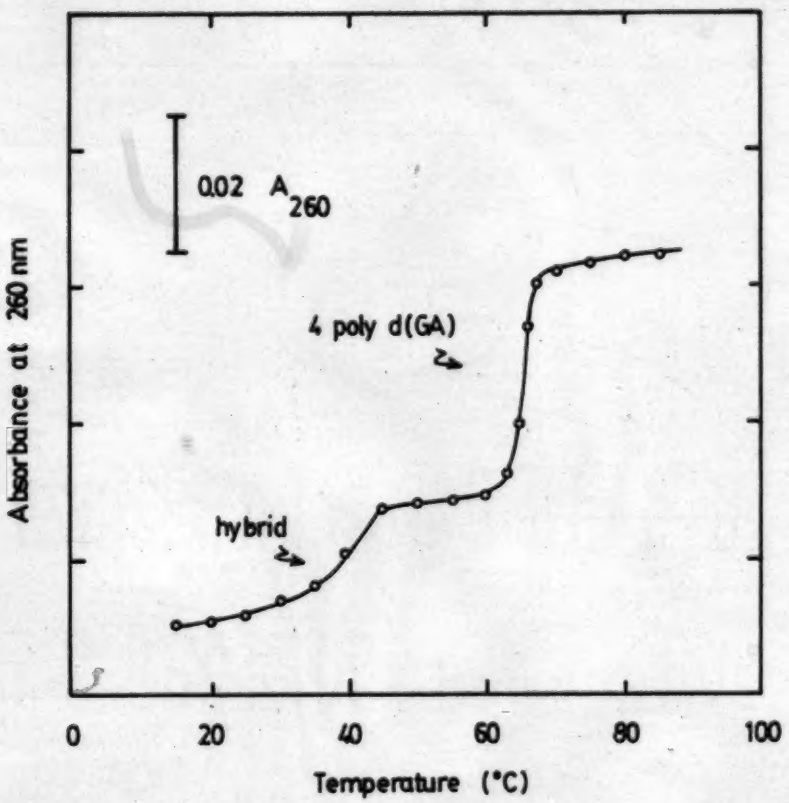


Figure 5-9 Melting of the poly d(GA)·poly d(TTC) hybrid in 0.40 M NaCl, TE. No transition is observed when only poly d(TTC) is heated. The upper transition corresponds to that expected of poly d(GA) alone. Initial absorbance was 0.330 A_{260}.

Clear evidence is presented here for the existance of extrahelical T's bracketed by CT or CC sequences. No good evidence of any other kind of extrahelical base has been seen in other DNA·DNA duplexes (see Chapter VI for uncertain structures). Thymine is the weakest stacking base (Solie and Schellman, 1968) and pyridones have the highest affinity for solvent water (Cullis and Wolfenden, 1981) two factors favouring an extrahelical conformation. Extrahelical A's and G's were observed in the RNA strand of DNA·RNA hybrids. Considering that analogous DNA·DNA hybrids either didn't form or had a different stoichiometry this suggests extrahelical bases are more favoured in RNA than in DNA. A more quantitative argument in agreement with this point can be developed by comparing the depression of T_m caused by extrahelical bases. For example the difference in T_m between poly d(TCC)·poly d(GGA) and the poly d(GGA)·poly d(TCTC) hybrid can be used to determine a destabilization index (Lomant and Fresco, 1972) D, which is a semiquantitative measure of the destabilization introduced into a helix by one extrahelical base per 10 base-pairs (Fink and Crothers, 1972)*. In 0.10 M NaCl D is 12° for the hybrid which is 5-9° higher than that of RNA duplexes containing small numbers of extrahelical uridines (Lomant and Fresco, 1972). This increased destabilization also suggests that extrahelical bases are less energetically favourable in DNA than they are in RNA. This difference in stability can be rationalized by model building. A good bridging hydrogen-bond can be formed between the 2'-hydroxyl and furanose oxygen of riboses bracketing an extrahelical residue.

* $D = [T_{mo} - T_{mx}]/[10x/(1-x)]$ where T_{mo} and T_{mx} are the T_m's of perfect and hybrid helices respectively and x is the fraction of extrahelical residues in one strand.

Although nearest neighbour and base-specific effects are probably important factors determining which bases will exist in an extrahelical configuration (Lomant and Fresco, 1975; Fink and Crothers, 1972; Wang and Kallenback, 1971) the data presented here does not permit further generalizations regarding these effects.

Conclusions

Extrahelical bases can exist in DNA duplexes containing defined base sequences. Such structures are less favoured in DNA than in RNA and in DNA·RNA hybrids occur in the RNA strand. Only thymidine has been observed in an extrahelical configuration in DNA duplexes. This is in contrast to RNA duplexes in which all four bases have been observed to exist extrahelically.

VI. Hybrid Polynucleotide Structures Containing Non-Watson-Crick Base-Pairs.

Introduction

Among the unusual structures that can be formed by polynucleotides there exists a class containing non-Watson-Crick base-pairs incorporated into otherwise normal duplexes. Table 6-1 summarizes most of the available papers describing such unusual pairing arrangements. These hybrids are distinguished from other types of non-classical polynucleotide structures in that the non-Watson-Crick base-pairs must be approximately isomorphous with ordinary base-pairs in order to coexist within Watson-Crick helices.

The evidence for non-classical base-pairing in random copolymers has been based primarily upon mixing curve analysis (ie, 1:1 strand stoichiometries are observed regardless of the fraction of non-complementary residues) while more recently NMR (eg, Romaniuk et al, 1979) and other methods (Dodgson and Wells, 1977) have also been used to probe such structures. In RNA the effects of such non-complementary pairs on helix structure are, reportedly, to greatly reduce the T_m and cooperative melting length (more so than extrahelical residues) while having little effect on the T_m dependence on salt concentration (Lomant and Fresco, 1975).

Many pairing schemes have been published which might account for these mismatches (Donohue and Trueblood, 1960; Crick, 1966; Drake and Baltz, 1976; Topal and Fresco, 1976a,b) the best known of which is the "wobble" hypothesis (Crick, 1966) which was originally proposed in order to rationalize flexibility in codon-anticodon pairing. All the

Table 6-1

Non-Watson-Crick Pairing in Hetropolymer Complexes

Random Copolymers

Polymers	f_{nc}	Comments	References
poly rA·poly r(U,I)	$f_I \leq 0.28$	I·A pairing	Wang & Kallenbach (1971)
poly rI·poly r(C,A)	$f_A \leq 0.30$	I·A pairing	Lomant & Fresco (1975)
poly rU·poly r(C,A)	$f_c \leq 0.30$	I·U pairing	Lomant & Fresco (1975)
poly rI·poly r(C,U)	$f_U \leq 0.33$	I·U pairing	Wang & Kallenbach (1971)
poly rU·poly r(G,A)	$f_G \leq 0.32$	G·U pairing	Lomant & Fresco (1975)
poly rC·poly r(I,A)	$f_A \leq 0.12$	A·C pairing	Lomant & Fresco (1975)

Defined Sequences

Polymers	Comments	References
poly dG·d($C_{12}A_xC_y$)	possibly G·A pairing	Dodgson & Wells (1977)
poly dG·d($C_{12}G_xC_y$)	possibly G·G pairing	Dodgson & Wells (1977)
r(CAGUG)·r(CAUUG)	G·U "wobble"	Romaniuk et al (1979)
2 poly d(GT)	G·T "wobble"	Early et al (1978)
2 poly r(GU)	G·U "wobble"	Gray & Ratliff (1977)
poly r(GGG)·poly d(TCC)	G·T "wobble"	Paetkau et al (1972)

Note: f_{nc} = fraction of non-complementary residues in the heteropolymer strand. In higher proportions these residues start becoming extrahelical.

wobble associations predicted by Crick have been seen in these hybrid structures (U·G, I·U and I·A) and so have some not considered likely (possibly G·G and G·A and certainly A·C)(Topal and Fresco, 1976a,b). Other types of oppositions appear to result in extrahelical structures (Chapter V) although the influence of neighbouring sequences can't be ignored.

The most well characterized of the non-Watson-Crick base-pairs are G·T(U) wobbles which have been studied by mixing curve analysis and UV spectroscopy (Lomant and Fresco, 1975), CD (Gray and Ratliff, 1977) and NMR (Romaniuk et al, 1979; Early et al, 1978). This pair has also been observed in the crystal structure of yeast $tRNA^{phe}$ (Quigley and Rich, 1976). NMR shows clearly that such pairs are locally unstable and have a T_m significantly lower than that of other base-pairs surrounding it (Romaniuk et al, 1979). No evidence of G^*·T or G·T^* tautomer pairs is seen. NMR spectra are reported to be consistent only with a wobble hydrogen bonding pattern (Early et al, 1978; Geerdes and Hilbers, 1979). The geometry of the other non-Watson-Grick base-pairs is still very poorly understood. They are all reported to be significantly less stable than G·T(U) pairs however.

In this chapter a number of hybrid structures are described which appear to contain further examples of unusual base-pairing.

Hybrids Containing G·T Pairs

Three pairs of polymers hybridized with a stoichiometry that strongly suggested G·T pairing was taking place. These are indicated in Table 6-2. The physical properties of these hybrids are summarized in Table 6-3. No attempt was made to measure melting slope because of the unknown contribution of length heterogeneity to transition width.

Table 6-2

Probable Structures of Hybrid Polymers Containing Non-Watson-Crick base-pairs.

Hybrid	Ratio of first to second strand	Structure
G·T Oppositions		
poly d(GAA)·poly d(TTT)	1:1.07	-GAAGAA- -TTTTTT-
poly d(GGA)·poly d(TTC)	1:1.04	-GGAGGA- -CTTCTT-
poly r(GAGA)·poly d(TTT)	1:0.78	-TTTTTT- -GAAGAA- G G
A·C Oppositions		
poly d(AAA)·poly d(TTC)	1:1.01	-AAAAAA- -TTCTTC-
poly d(GGA)·poly d(CCC)	1:1.04	-GGAGGA- -CCCCCC-
T·C Oppositions		
poly d(TG)·poly d(CC)	1:1.01	-TGTGTG- -CCCCCC-

Note: Strand ratios were determined as indicated in the legend to Figure 5-1(b).

Table 6-3

Melting Points of Hybrid Polymers Containing Non-Watson-Crick Base-Pairs in 0.1 M NaCl, TE Determined by UV Spectroscopy.

Hybrid	T_m (C°)	dT_m/d log[NaCl] (C°)	D_{GC}[1] (C°)	D_{AT}[1] (C°)
G·T Oppositions				
poly d(GAA)·poly d(TTT)	27.5	23	9.6	8.4
poly d(GGA)·poly d(TTC)	43.5	14.2	9.1	7.7
poly r(GAGA)·poly d(TTT)	32	10	---	---
A·C Oppositions				
poly d(AAA)·poly d(TTC)	27	28	9.7	8.5
poly d(GGA)·poly d(CCC)	45	15	11	8.8
T·C Oppositions				
poly d(TG)·poly d(CC)	14[2]	21	8.6	7.3

Note: 1. D is calculated as described earlier using reference data compiled in Table 5-3. D_{GC} and D_{AT} are destabilization indices calculated relative to perfect duplexes containing G·C or A·T transitions (respectively) in place of the non-Watson-Crick pair.

2. Extrapolated from T_m's at higher salt concentrations.

The poly d(GAA)·poly d(TTT) hybrid annealed with a 1.07:1 ratio sufficiently close to 1:1 to suggest G·T pairing. No peculiar physical properties were noted.

The poly d(GGA)·poly d(TTC) hybrid was more interesting because it was apparent that more than one structure could be formed by these two strands. Figure 6-1 illustrates the effect of salt concentration on the melting point of this hybrid. Above 0.41 M NaCl the single transition associated with hybrid melting became two transitions. The tetraplex transition became smaller as the salt concentration increased and disappered, also at 0.41 M NaCl. Controls showed that even in the absence of polypyrimidine strands the tetraplex transition still disappeared above 0.41 M NaCl. Figure 6-2 shows the two hybrid transitions in 1M NaCl. By varying the amount of polypyrimidine it became clear that each complex was composed of different strand ratios. Based on a UV mixing curve the stoichiometry of the lower melting hybrid is probably 1.5:1 and the second hybrid (from Figure 6-2) probably 1:1. Most likely both the single low salt and more stable high salt form contain G·T pairs while the less stable high salt form contains extrahelical bases. Its not clear why it should be destablized by increasing salt concentration. This is a phenomenon characteristic of charged helices (eg. Vetterl and Guschelbauer, 1979)but changing the pH from 8.0 to 7.0 had no significant effect on T_m. The disappearance of the poly d(GGA) tetraplex transition wasn't an isosbestic point artifact and it seems more than a coincidence that the critical salt concentrations are identical in both cases. These peculiar observations have to remain unexplained at present.

The poly r(GAGA)·poly d(TTT) hybrid has been characterized

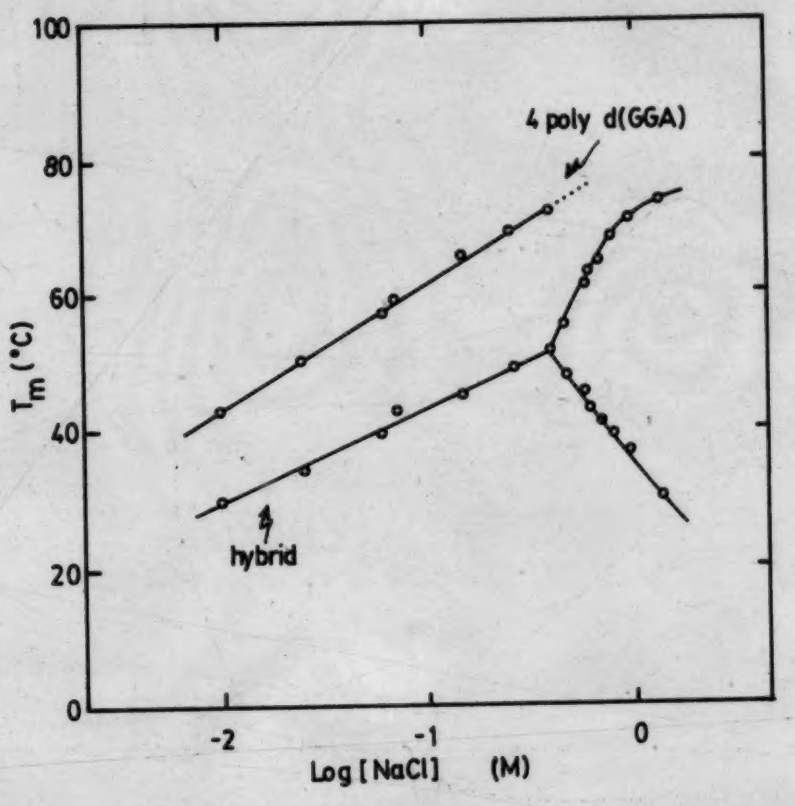


Figure 6-1 The effect of NaCl concentration on poly d(GGA)·poly d(TTC) melting. Poly d(GGA) and poly d(TTC) were mixed in 1:1 strand ratio in the presence of the indicated concentration of salt, allowed to anneal (typically 30min.) and the T_m determined by following the absorbance as a function of temperature at 260 nm. Buffer was TE.

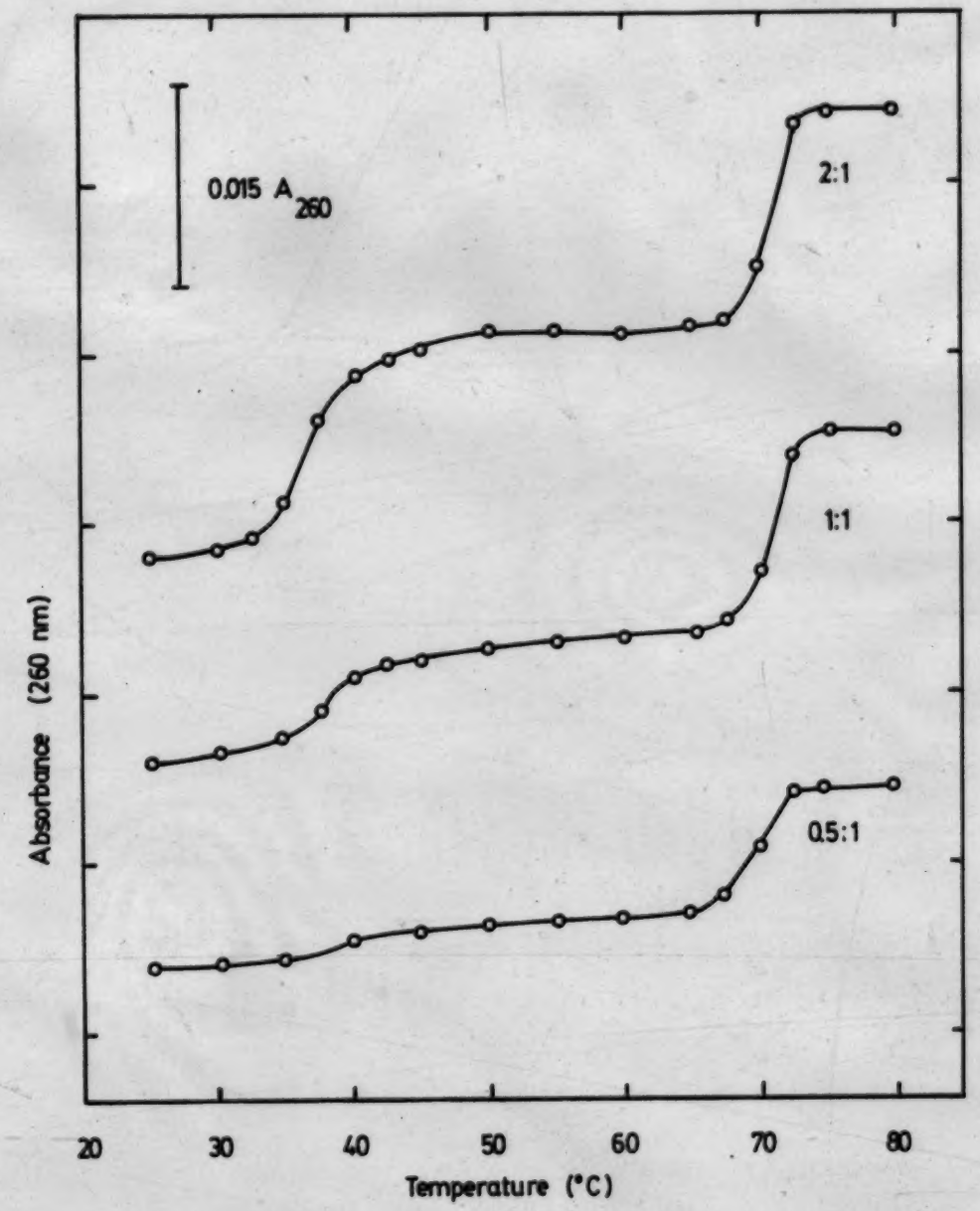


Figure 6-2 Effect of strand ratio on relative hyperchromicities in 1.0 M NaCl, TE. Samples contained 10.7 µM poly d(GGA) and 5.4 (0.5:1), 10.8 (1:1) or 21.7 (2:1) µM poly d(TTC) in 300 µl. ΔA_{260} was followed as a function of temperature.

(Chapter V) as probably containing half the G's extrahelical and half base-paired.

Hybrids Containing A·C Pairs

Two hybrids were prepared that clearly appeared to contain A·C oppositions. Poly d(A) and poly d(TTC) annealed in a 1:1 ratio by fluorescence (Table 6-2). As expected considering that the Watson-Crick pairs are A·T the T_m was low but showed a surprisingly large sensitivity to salt concentration (Table 6-3). Because this sensitivity is sometimes characteristic of triplexes and because a triplex may not have been detected by ethidium bromide fluorescence measurements a UV mixing curve was also constructed. This gave a stoichiometry identical to that observed by fluorescence (Figure 6-3).

Inspection suggested that poly d(GGA) should be able to anneal to poly d(C) but this hybrid proved difficult to prepare reproducibly by the usual hybridization method in which heat denatured polypurine was added to the polypyrimidine strand, salt added, and the sample incubated at room temperature. It could be formed reproducibly by heat denaturing a mixture of the two strands in alkali, neutralizing and allowing to anneal. The melting transition of this hybrid showed the same peculiarity associated with poly d(GGA) melting. As the salt was increased the hyperchromicity progressively decreased until above 0.19 M salt no transition could be seen at 260 nm (Figure 6-4). This wasn't an isobestic point artifact. Although an optical transition cannot be seen in 0.5 M NaCl the strands still appeared to be hybridizing as judged by ethidium bromide fluorescence. These curious observations are discussed in more detail later. Once formed the hybrid was very stable (Table 6-3).

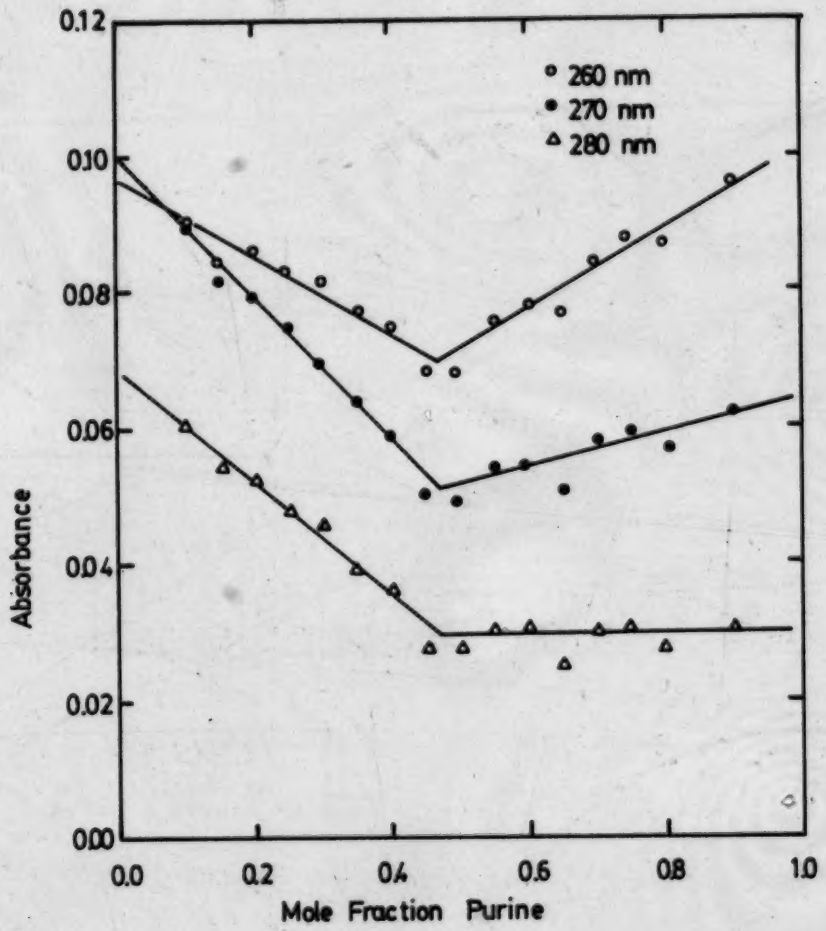


Figure 6-3 UV mixing curve·poly dA and poly d(TTC). Samples (300 µl) contained 13.3 µM total polynucleotide in 0.5 M NaCl, TE. After annealing one hour each sample was transfered to a microcuvette centrifuged and the absorbance recorded. Results are taken from two separate experiments.

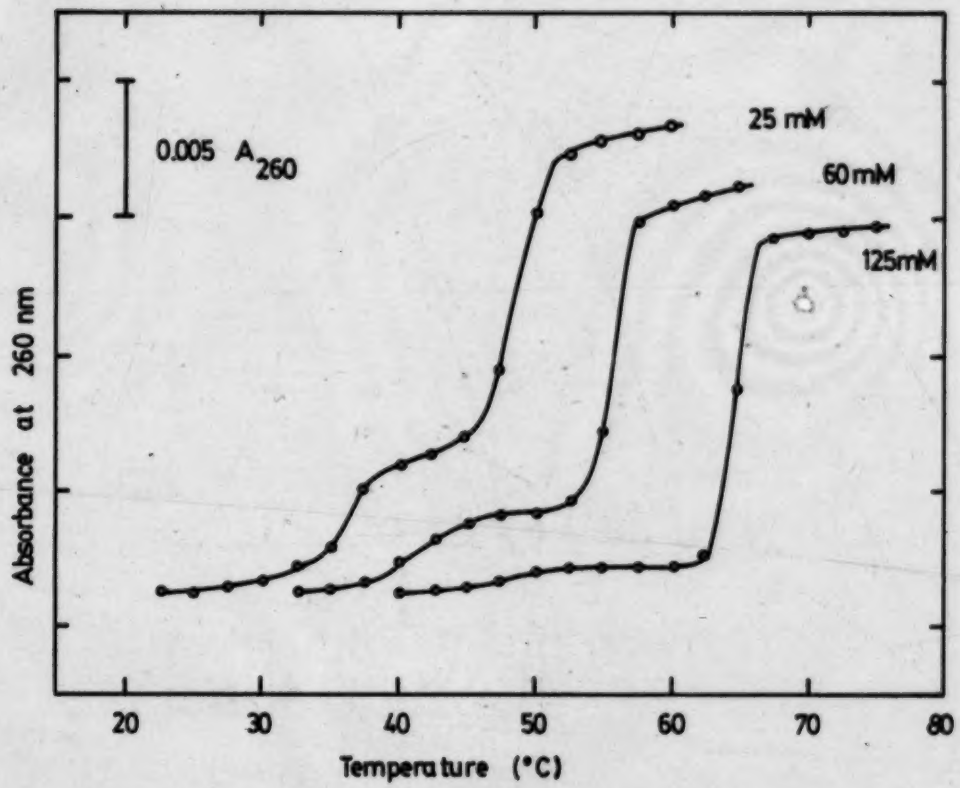


Figure 6-4 Melting of the poly d(GGA)·poly d(CCC) hybrid in different concentrations of salt. Poly d(GGA) was hybridized to poly dC as described in the text and desalted into 50 mM NaCl, TE. Samples for T_m measurements were then diluted into 250 µl along with the salt required to give the final indicated NaCl concentrations. Initial absorbance was 0.21 A_{260}. ΔA_{260} was followed as a function of temperature.

Hybrids Containing Pyrimidine·Pyrimidine Oppositions

Poly d(TG) and dC annealed with a 1:1 stoichiometry suggesting a hybrid containing T·C oppositions could form (Table 6-2). As one would expect considering the high frequency of helix defects the hybrid was exceedingly unstable. Poly d(TG) also annealed to poly d(TC) and although the stoichiometry was not determined it is probably also 1:1. The poly d(TG)·poly d(CT) hybrid was slightly less stable than the poly d(TG)·poly d(CC) hybrid (25° vs 28.5° in 0.5 M NaCl, TE respectively). No evidence of a poly d(GT)·poly d(GT) hybrid was detected either by fluorescence or absorbance methods as expected since the T_m of this hybrid appears to be less than 20° in 0.5 M NaCl (Early et al, 1978; Gray and Ratliff, 1977).

Other Hybrids of Uncertain Structure

Table 6-4 summarizes the strands which hybridized in a stoichiometry which was difficult to equate with any particular structure. In all cases structures containing blocks of three Watson-Crick pairs followed by three non-Watson-Crick (G·T and A·C) pairs can be drawn but there is no experimental data to support these over structures containing extrahelical bases in both strands.

Strands Which Do Not Anneal

It is also instructive to consider hybrids which apparently do not form since this illustrates the limits to which defects can be accommodated into helices (Table 6-5). A trace of fluorescent product was observed when poly d(GAA) and poly d(TCC) were hybridized but hybridization didn't saturate in the presence of excess poly d(TCC).

Discussion

A number of hybrids have been characterized as containing A·C,

Table 6-4

Hybrid Polymers of Uncertain Structure

Hybrid	Ratio (pyr:pur)	T_m (C°)	dT_m/d log [NaCl]
poly d(GAAGAA)·poly d(TCTCTC)	1.02:1	27	32
poly d(GAGAGA)·poly d(TCCTCC)	1.09:1	45	21
poly r(GAGAGA)·poly d(TTCTTC)	1.04:1	46	17

Note: Strand ratios were determined as indicated in the legend to Figure 6-1(b). T_m is in 0.10 M NaCl, TE.

Table 6-5

Strands Which Failed to Hybridize in 0.5 M NaCl, 0.10 M Tris pH 8 as Determined by Ethidium Bromide Fluorescence.

poly dA + poly d(TC) , poly d(TCC), poly d(TG)

poly d(GAA) + poly d(TCC) , poly dC

poly d(GA) + poly dC

poly d(GGA) + poly dT

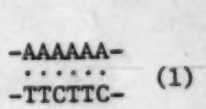
G·T and pyrimidine·pyrimidine oppositions. As expected such oppositions destablize relative to perfect Watson-Crick helices but not to as great an extent as was observed in hybrids containing extrahelical bases (D = 6.5 - 11° vs 11-12°). This is not unexpected as it seems reasonable that the amount of backbone distortion associated with an extrahelical base is considerably greater than that required to incorporate G·T or A·C pairs into a regular helix. This observation, however, is at variance with statements that intrahelical non-Watson-Crick pairs are more destabilizing than are extrahelical bases (Lomant and Fresco, 1975). Unfortunately these claims are based to a great extent on unpublished data. Clustering of some residues in polynucleotide phosphorylase synthesized RNA's has been reported (Greene et al, 1978) so it may be that these workers are describing the properties of non-complementary base runs rather than single non-complementary oppositions. RNA may also behave differently than DNA.

In general $dT_m/d \log[NaCl]$ values are significantly higher in defective helices than in perfect helices (median 21°, range 10-32° vs median 14°, range 10-16°). Despite claims to the contrary (Lomant and Fresco, 1975) this does not appear to be a reliable way to differentiate helices containing extrahelical bases from those containing non-Watson-Crick base-pairs. Clearly the problem with this measurement is that it reflects both overall helix geometry as well as the geometry of local distortions. Since one doesn't know the relative contribution of each to the salt dependency one is unlikely to be able to discriminate between the different structures with this technique.

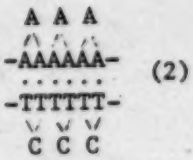
Neither G·T, A·C or pyrimidine·pyrimidine oppositions showed any significant differences in destabilizing tendencies. A tentative

conclusion is that such oppositions destablize primarily by disrupting stacking and backbone geometry and that if they are paired the hydrogen bonding provides little compensatory stabilization. Here, as elsewhere, one has to be careful not to overinterpret the significance of D since it is only a semiquantitative parameter lacking a really firm theoretical foundation (Fink and Crothers, 1972).

Whether the oppositions described here are actually base-paired is uncertain since they have only been characterized as to strand stoichiometry. For example it is possible that instead of (1) the structure is (2)



(1)

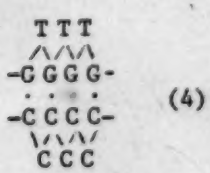


(2)

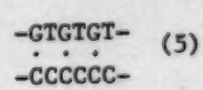
but then it is not obvious why the structure isn't (3)

(3)

Conceivably (3) will be less favoured than (2) if one postulates that extrahelical bases in opposition are more stable than are extrahelical bases adjacent to a perfect strand. While it is probable that G·T and A·C do base-pair it seems very unlikely that pyrimidine·pyrimidine pairing occurs in the poly d(GT)·poly d(CC) hybrid. Thus the structure is more likely (4) than (5).



(4)



(5)

Four is especially favoured since it makes possible the stacking of G on G. It should be noted that to date there exists no evidence of

pyrimidine·pyrimidine pairing in Watson-Crick helices (Topal and Fresco, 1976b; Lagerkvist, 1981) although the mitochondrial genetic code may be read by a process involving U·U or U·C recognition at the third codon position (Lagerkvist, 1981).

It is curious that a number of oppositions that "sensibly" should result in extrahelical structures don't. For example poly d(GA) and poly d(TCC) annealed 1:1 and while the structure is uncertain it clearly isn't (6).

```
-GAGAGA-
 ......
-CTCTCT-     (6)
  \/ \/ \/
   C C C
```

which would have a 1.33:1 pyrimidine to purine ratio. This structure is identical to that formed by poly d(TTC) and poly d(GA) except that it contains extrahelical C's instead of T's. Thus in the hybrids studied here T seems to exist extrahelically while C doesn't. This has an interesting implication concerning the origin of frameshifting mutations since it implies that spontaneous frameshifts should occur with higher frequencies in regions where T's can loop out than where C's can. Spontaneous frameshift mutations do occur with high frequency at sites containing stretches of A (or T) (Streisinger et al, 1966; Pribnow et al, 1981). The three "hottest" sites in the rII gene each contain seven contiguous A's for example (Pribnow et al, 1981). Unfortunately comparable stretches of G and C don't exist in these sequences so it is not clear whether these high mutation frequencies are peculiar only to runs of A. Drug induced frameshift mutations are known to originate in a variety of sequences (Drake, 1970; Ames, Lee and Durston, 1973) but this is more likely a reflection of binding specificity.

How can one rationalize the peculiar properties of the poly d(GGA)·poly d(CCC) hybrid? Unlike the other hybrids this is the only one in which both strands can form ordered structures after the hybrid melts. Poly d(GGA) forms a tetraplex (Chapter IV) and poly dC forms a stacked single stranded helix like that formed by poly rC (Arnott et al, 1976). Formation of the hybrid must disrupt both structures and may be a path dependent process as has been observed in other polynucleotide complexes (Inman, 1964). Since new structures form on melting it is not surprising that the transition disappears above a certain salt concentration. Presumably any hyperchromicity associated with hybrid melting is matched by hypochromicity associated with tetraplex and poly dC helix formation and the transition vanishes.

Finally why didn't some polymers hybridize? Poly d(GGA) and poly d(TCC) hybridized weakly but not well enough to determine the stoichiometry. In all other cases hybridization would produce a helix with a high frequency of mismatches (1:2 or more) and as the poly d(TG) poly d(CC) hybrid T_m indicates these are not likely to form under these conditions.

Conclusions

A number of non-complementary polymers hybridize with a stoichiometry that suggests G·T and A·C pairing is taking place. Other hybrids contain pyrimidine-pyrimidine oppositions but these are unlikely to be base-paired. These hybrids are less stable than perfect helices and their T_m's are more sensitive to salt. Hybrids containing extrahelical bases cannot be distinguished from those containing non-complementary oppositions on the basis of destabilization indices or salt dependency. At least one of these hybrids forms different structures depending on conditions.

VII. Physical Studies

Having identified a number of unusual hybrid polynucleotides it is worth considering some of their physical properties.

Buoyant Density Studies

Synthetic polyribonucleotides have long been known to alter the buoyant density of single-stranded DNA's in CsCl (Kubinski *et al*, 1966) and this technique has been used to separate the complementary strands of denatured DNA's in both CsCl and $CsSO_4$ (Szybalski, 1968). This was at one time assumed to be due to the formation of hybrids over short regions of homology. How short these regions of homology need be was demonstrated by Paetkau *et al* (1972) who showed that a poly r(GGG)·poly d(TCC) complex could be prepared by RNA polymerase transcription of poly d(TCC) and had a buoyant density in $CsSO_4$ significantly different from both poly d(TCC) and poly r(GA). Similarly DNA·DNA hybrid formation can also be demonstrated by this technique and here the homology is even smaller. Typically free single strands band poorly at distinctly different densities (Figure 7-1(a)) while duplexes form sharp overlapping peaks (Figure 7-1(d)). Hybrids also banded as sharp overlapping peaks (Figure 7-1(b,c)). Thus the technique described by Szybalski does not simply rely on the formation of hybrids between homologous sequences.

It is possible to alter the buoyant density of λDNA by heat denaturing it in the presence of poly d(TCC) but the shift is small [*circa* 13 mg/ml with 1:1 poly d(TCC) to λDNA] and the technique isn't of any practical use.

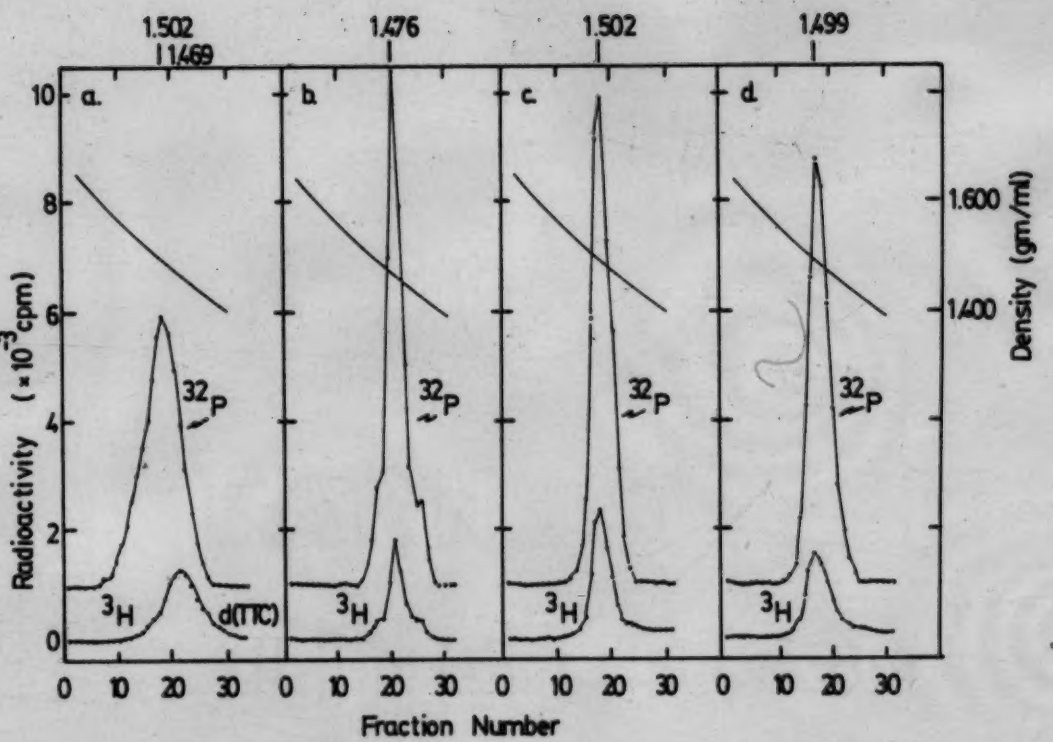


Figure 7-1 Buoyant density shifts accompanying hybrid formation. Preparative $CsSO_4$ gradients contained ^{32}P labelled poly d(GGA) and ^{3}H labelled polypyrimidine. Panel (a) single stranded polymers banded in separate but parallel tubes. Panel (b) poly d(GGA)·poly d(TTC). Panel (c) poly d(GGA)·poly d(TCTC). Panel (d) poly d(GGA)·poly d(TCC). Hybrids [panels (b) and (c)] band sharply at intermediate densities as do proper duplexes [panel (d)].

Nuclear Magnetic Resonance

NMR spectroscopy is becoming increasingly popular as a technique for investigating polynucleotide structure (Gorenstein, 1981; Reid, 1981). ^{31}P NMR shows particular promise since ^{31}P chemical shifts are reported to be sensitive to bond and torsion angle effects (Haasnoot and Altona, 1979; Gorenstein, 1978). In yeast $tRNA^{phe}$ for example the bulk of resonances fall between -0.5 and -2.0 ppm (relative to 85% H_3PO_4) with a few others scattered over 1.5 to -5.0 ppm (Gorenstein et al, 1981). The identity of most of these scattered resonances is unknown and because of inadequacies in chemical shift theory they will have to be identified by empirical methods. Considering the odd phosphate conformation associated with an extrahelical base the NMR spectrum of poly d(GGA)·poly d(TCTC) was examined in the hope that such phosphates might be identified by a resonance of unusual chemical shift. Figure 7-2 shows these spectra.

Under conditions where the hybrid is quite stable (0.1 M NaCl, 20° below T_m) one peak centred at -0.5 ppm was observed. In 30 mM NaCl (10° below T_m) this peak became sharper and close inspection showed it to be two peaks separated by 0.14 ppm at -0.43 and -0.57 ppm. These results are very similar to those reported by Cohen et al (1981). In alternating copolymers such as poly d(AT) and poly d(IC) doublet resonances separated by 0.1-0.2 ppm were observed which upon further cooling coalesced into broad singlets. This was taken to be evidence of an alternating B-DNA structure (Klug et al, 1979). Thus the data in Figure 7-2 suggests there are at least two distinct phosphodiester geometries in poly d(GGA)·poly d(TCTC). These could also be explained as an equilibrium mixture of two interconverting conformations (Selsing et al,1978)

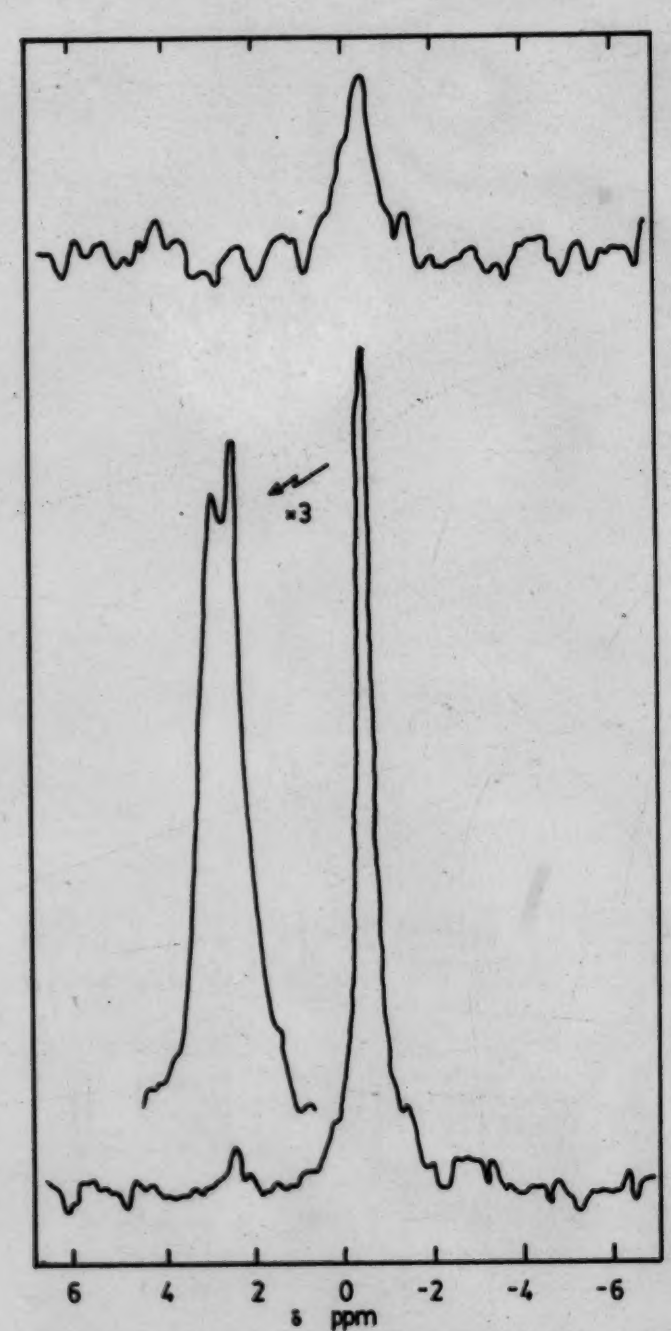


Figure 7-2 NMR spectra of poly d(GGA)·poly d(TCTC). Upper spectrum 150 μM poly d(GGA), 210 μM poly d(TC), 100 mM NaCl, TE and 33% D_2O in 1.5 ml. Lower spectrum 160 μM poly d(GGA), 220 μM poly d(TC), 30 mM NaCl, TE and 50% D_2O in 2.0 ml.

50,000 free induction decays each at 300 K° referenced to 85% H_3PO_4.

but if so it is slow on the time-scale of these experiments (less than 100 sec^{-1}) (Cohen et al, 1981; Gorenstein, 1981). Whether the "peak" at +2.5 ppm is real is debatable but in any case the intensity appears to be too small for it to be one of the hybrid phosphates. The failure to observe an unusual phosphorous resonance that might be characteristic of extrahelical bases and insensitivity problems (50,000 free induction decays takes about a day) caused further experiments to be abandoned.

Calorimetry

In quantitative terms how stable are the structures described in Chapters V and VI? T_m's provide only a relative estimate of stability and can give a distorted impression of what structure is likely to form under conditions where several are possible. High sensitivity differential scanning calorimetry (DSC) provides a direct method by which it is possible to determine the true stability of these structures.

Enthalpies of helix formation can be measured in a number of ways. From theoretical considerations van't Hoff enthalpies (ΔH_{VH}) can be derived by spectrophotometric methods from either the effect of concentration on melting temperature or from the shape of the melting transition (Martin et al, 1971; Fink and Crothers, 1972; Breslauer et al, 1975). Unfortunately the resulting enthalpies tend to be in some disagreement with each other (Martin et al, 1971) and with calorimetric data (Breslauer et al, 1975). The discrepancy has been attributed to base-line problems (Albergo et al, 1981) but the inherent assumptions that melting is a simple two-state process and that all thermodynamic transitions are spectrophotometrically detectable are also debatable. Because of this the thermodynamic properties of polynucleotides are

best derived by direct calorimetric methods thereby avoiding these uncertainties and assumptions.

Table 7-1 summarizes some of the thermodynamic properties of oligo and polynucleotides derived by high sensitivity calorimetry. The data on this subject is rather confusing since different experimental conditions make comparisons very difficult. It is also not clear from the literature whether the enthalpy of helix formation is temperature dependent or not. The distinction is important since it determines whether enthalpies can be compared at any temperature. Breslauer et al (1975) failed to observe a significant change in heat capacity before and after melting (at constant pressure $C_p = \Delta H/\Delta T$ so a temperature dependent enthalpy is revealed by a change in heat capacity during melting) while Filimonov and Privalov (1978) reported clear evidence of such effects. Because of this uncertainty and because comparisons with the literature are difficult it was also necessary in these studies to examine the calorimetric properties of the pyrimidine·purine DNA's from which the hybrid polymers derive.

Calorimetric Profiles

Figure 7-3 shows a typical calorimetry scan of a sample of poly d(TC)·poly d(GA) illustrating several interesting features. The small distortions in the base-line just prior to melting weren't reproducibly observed. They probably correspond to the premelting transitions sometimes observed by spectrophotometric methods (Cantor and Schimmel, 1981) but were not investigated further. Note also the effect of scan rate. The first melting was recorded at 1.063 deg·min^{-1} while the second was recorded at 0.899 deg·min^{-1}. The better base-line of the second melt was also a general feature of these experiments.

Table 7-1

Apparent Thermodynamic Properties of Helix Formation at $T = T_m$ Determined by Calorimetry.

Oligo or Polynucleotide	T_m (C°)	ΔH(Kcal·mol bp^{-1})	ΔS(cal·mol $bp^{-1}\cdot deg^{-1}$)	Reference
$d(GC)_3$				Albergo et al
1000 mM NaCl	67	-9.93	-27.4	(1981)
$r(A_7U_7)_2$				Breslauer et al
1000 mM NaCl	46	-7.1	-20.2	(1975)
poly rA·poly rU				Filimonov & Privalov
10 mM NaCl	37	-6.8	-22	(1978)
50 mM NaCl	50.5	-7.9	-24	
100 mM NaCl	57.5	-8.4	-25	
poly rA·poly rU				Newman & Ackerman
60 mM NaCl	49	-6.8	-21	(1967)
poly rA·poly rU				Krakauer & Stur-
18 mM NaCl	44.5	-7.38	-23.2	tevant (1968)

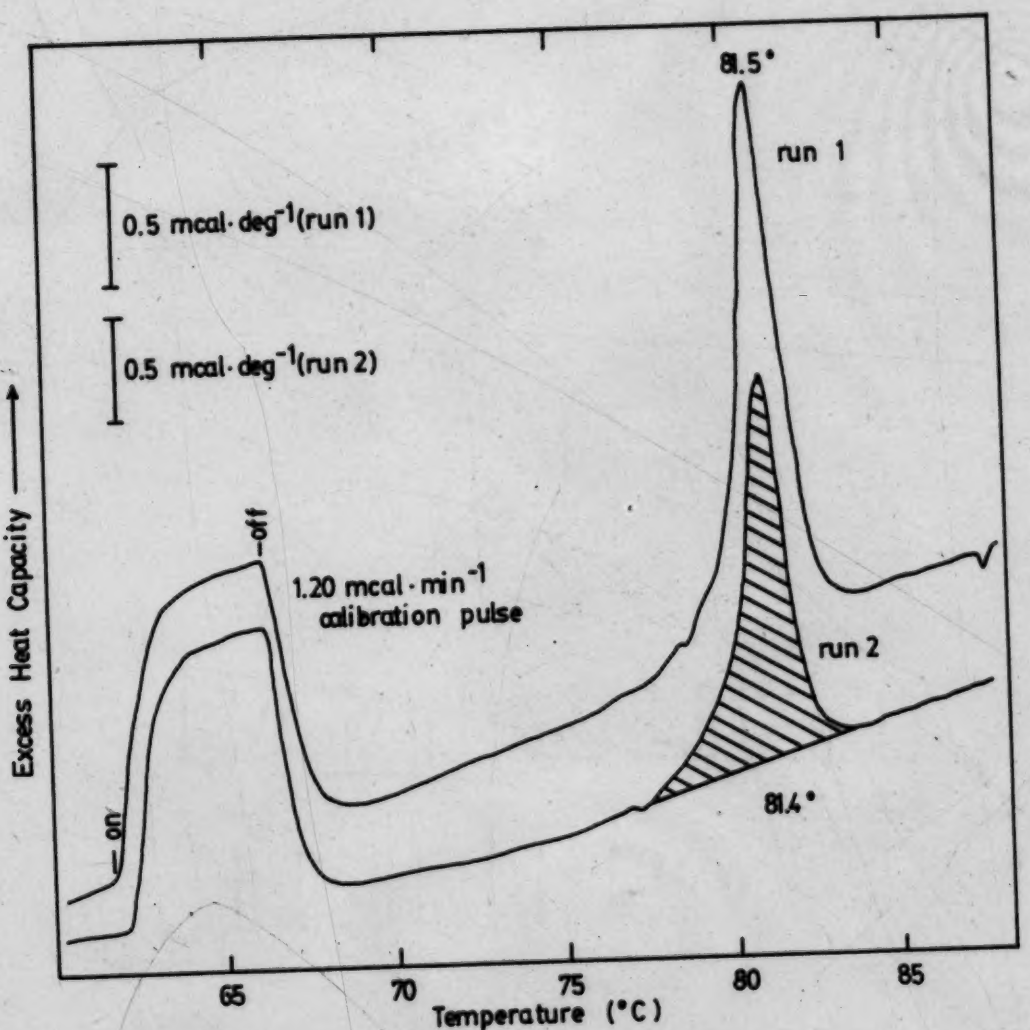


Figure 7-3 Melting of poly d(TC)·poly d(GA) followed by high sensitivity DSC. Sample contained 945 μM poly d(TC)·poly d(GA), 0.10 M NaCl, TE in 0.876 ml. After Run 1 the sample was allowed to cool slowly overnight and Run 2 recorded next morning.

A temperature dependent enthalpy is detectable by an abrupt change in base-line slope centred at $T = T_m$. Within the uncertainties associated with base-line drift this wasn't seen (eg. Figure 7-3). The only possible exception to this was poly d(G)·poly d(C) where the size of the dependence of enthalpy on temperature was of a similar magnitude to that reported by Filimonov and Privalov (1978) circa 30 mcal·mol bp^{-1}·deg^{-1}. For the purposes of discussion it will be assumed that melting enthalpies are temperature independent in all other cases. See Appendix 2.

Calculations

For a general discussion of DSC see Mabrey and Sturtevant (1978). Calorimetric enthalpies are determined by direct integration of the excess heat capacity curve. This area is then converted to mcal by comparison with a calibration pulse of known heat input rate. The curve area indicated in Figure 7-3 was measured to be 11.8(3) cm^2 and the calibration pulse height 6.0(4) cm. The calibration pulse corresponds to a heat input rate of 1.20 mcal·min^{-1}, the heating rate was 0.89(9) deg·min^{-1} and the pen rate 1.21(9) deg·cm^{-1}. From this it can be calculated that the calibration is 0.21(6) mcal·cm^{-2} and the peak area 2.5(8) mcal. The sample contained 0.78 μmol (Pi) polynucleotide so the transition enthalpy was 3.2(9) Kcal·mol Pi^{-1} or 6.5(9) Kcal·mol bp^{-1}.

The van't Hoff enthalpy (ΔH_{VH}) can also be determined from these traces. This is of interest since the ratio $\Delta H_{VH}/\Delta H_{cal}$ indicates the size of the cooperative unit (Mabrey and Sturtevant, 1978). Within ±10% ΔH_{VH} can be conveniently calculated by the formula

$$\Delta H_{VH} \cong 6.9 \frac{T_m^2}{\Delta T_{1/2}}$$

where T_m is the midpoint of the transition (the transition maximum within ±1 deg) in K° and $\Delta T_{1/2}$ the width of the transition in degrees.

At $T = T_m$ the entropy change is simply defined by

$$\Delta S = \frac{\Delta H}{T_m}$$

where T_m is again in K°. ΔS is a salt dependent function (Record et al, 1981) so 0.10 M NaCl, TE was chosen to be the standard conditions for which it was calculated.

Assuming ΔH and ΔS to be temperature independent functions (certainly true within the data error) the free energy is by definition

$$\Delta G = \Delta H - T\Delta S$$

and from this K_{eq} is defined by

$$K_{eq} = \exp(-\Delta G/RT)$$

It is also theoretically possible to derive from thermodynamic parameters the average charge spacing along the helix axis. In Appendix 1 it is demonstrated that this relationship is given by :

$$b_{helix} = 4.3 - \frac{3.12\ \Delta H}{T_m^2} \cdot \frac{d\ T_m}{d\ \log[NaCl]} \quad (\text{Å})$$

ΔH is the enthalpy per residue of helix melted, T_m the melting temperature (in K°) and $d\,T_m / d \log [NaCl]$ the slope relating melting point and salt concentration.

Results - Ordinary DNA Duplexes

Sensitivity is directly proportional to heating rate in DSC but too high a heating rate can introduce disequilibria artifacts. The effect of different heating rates was examined in order to determine the rate providing maximal sensitivity without heating artifacts. It was

found that heating rates of 1 deg·min^{-1} or less were satisfactory and were used in all subsequent experiments. This heating rate may not be appropriate for polymers longer than those used in these experiments.

Figure 7-4 illustrates the range of errors associated with repeated measurements. Commercial polynucleotides must be used with caution since their history is rather uncertain. This was especially true of poly dG·poly dC since unlike the other polymers it doesn't reanneal properly on slow cooling overnight and has been observed to form a variety of complexes of unknown structure on freezing. Thus care was taken to avoid freezing or denaturation prior to calorimetric measurements. Figure 7-4 also illustrates the observation (confirmed with other polymers) that within the error of these measurements melting enthalpies didn't vary significantly with salt concentration.

Melting enthalpies for the five different pyrimidine·purine DNA's are shown in Figure 7-5 and summarized in Table 7-2. Average values for enthalpies of A·T and G·C pair formation were calculated by assuming the contribution to overall stability to be simply proportional to base composition, cross-substituting in all possible ways the five equations in two unknowns and averaging the results. Its difficult to say whether the observed enthalpies deviate significantly from those that can be calculated using a simple average. The results presented in Table 7-2 compare reasonably with previously published values (Table 7-1) although the enthalpy of poly dG·poly dC helix formation does appear to be significantly lower than $d(GC)_3$. These enthalpies are significantly lower than those calculated by Borer *et al* (1974) from van't Hoff analysis of oligoribonucleotide melting. The cooperativity varied from experiment to experiment and no obvious correlation between cooperative

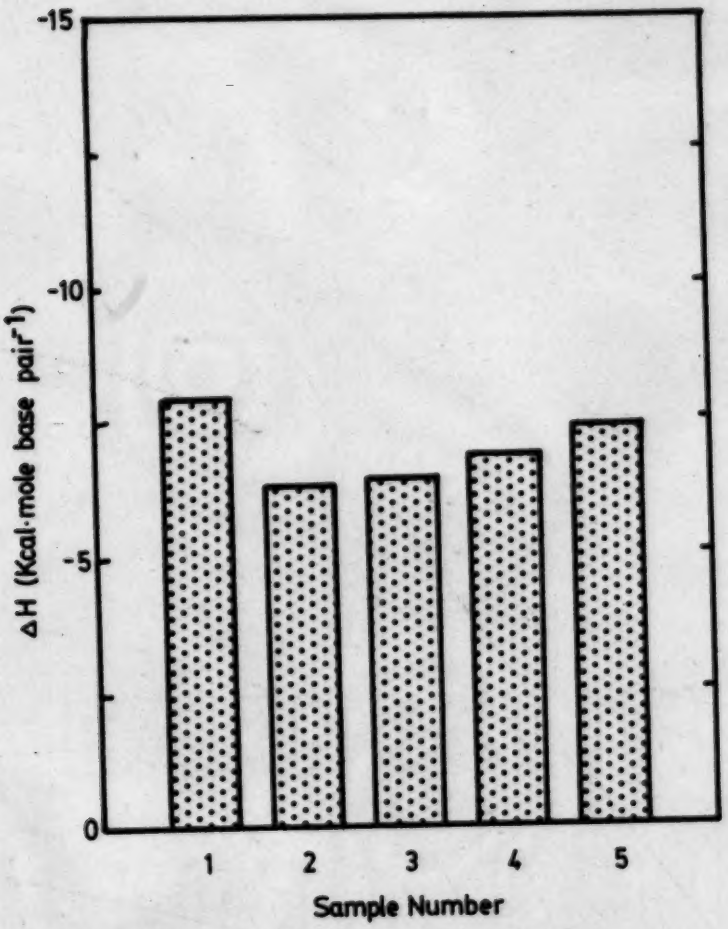


Figure 7-4 Errors associated with replicate calorimetric measurements. Enthalpy of poly dA·poly dT helix formation. (1) Commercial polynucleotide in 5 mM NaCl, TE, Run 1. (2) Sample from (1) cooled slowly overnight and remelted. (3) Laboratory sample in 5 mM NaCl, TE, Run 1. (4) Sample from (3) cooled overnight and remelted. (5) Laboratory sample in 0.10 M NaCl, TE.

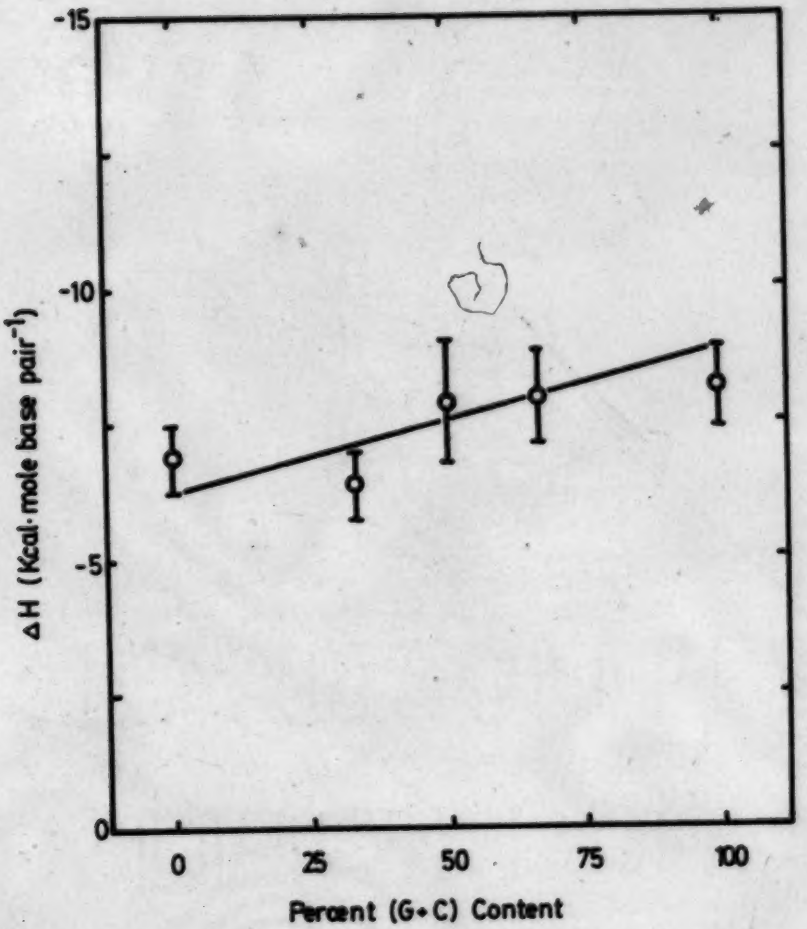


Figure 7-5 Enthalpies of helix formation in pyrimidine-purine DNA's. Error bars show average error around the mean. Line calculated from average enthalpies of A·T and G·C base formation. See Footnote to Table 7-2.

Table 7-2

Apparent Thermodynamics of Helix Formation in 0.10 M NaCl, TE.

Polymer	T_m(K°)	ΔH(Kcal·mol bp^{-1})	ΔS(cal·mol bp^{-1}·deg^{-1})
poly dA·poly dT	342.5	-6.9 ± 10%	-20 ± 10%
poly d(TTC)·poly d(GAA)	348.5	-6.4	-18
poly d(TC)·poly d(GA)	355	-7.9	-22
poly d(TCC)·poly d(GGA)	361	-8.0	-22
poly dG·poly dC	372.5	-8.2	-22
average A·T		-6.27	-18.7
average G·C		-8.76	-23.8

Note: Except for poly dG·poly dC enthalpies were determined in both 0.005 and 0.10 M NaCl, TE and are average of four or five measurements. Poly dG·poly dC melting was measured only in 0.005 M NaCl, TE and two measurements were made. Errors refer to average deviations from the mean (Skoog and West, 1976) and is the overall average error for all determinations.

length and polymer identity or polymer length was noted. For all experiments $\Delta H_{VH}/\Delta H_{cal}$ averaged to circa 80 ± 50.

Results - Hybrid Polynucleotides

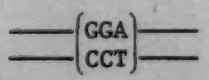
What are the relevant thermodynamic properties of hybrids containing helix defects? This data is summarized in Table 7-3. The hybrids melt with two transitions the first being the hybrid and the second the tetraplex. The total hybrid enthaply of formation from single strands is the sum of these two enthalpies. Since an extrahelical base contributes nothing to helix stability a better estimate of stability comes from reporting the poly d(GGA)·poly d(TCTC) hybrid enthalpy on a per base-pair basis and ignoring the extrahelical base (Kcal·mol Pi^{-1} x 7/3). The other enthalpies are reported assuming the mismatches are base-pairs. The larger proportionate errors are due to there being two transitions and difficulties in preparing large amounts of hybrids.

Within these limits it is quite clear that helix defects destabilize ΔH relative to the perfect helices. In poly d(GGA)·poly d(TCTC) this is one extrahelical base per three base-pairs so:

$\Delta H = 3 \times -4.4(8)$

$= -13.(4)$ Kcal·mol^{-1}

$\Delta S = -42$ cal·mol^{-1}·deg^{-1}



$\Delta H = 3 \times -8.0(2)$

$= -24.(1)$ Kcal·mol^{-1}

$\Delta S = -66$ cal·mol^{-1}·deg^{-1}

On a per defect basis an extrahelical T is unfavoured enthalpically by circa 10 Kcal·mol^{-1} which about corresponds to the loss of a complete G·C pair. At 25° this corresponds to a destabilizing free energy $\Delta G^{\circ} \approx 3.5$ Kcal·mol^{-1} which is somewhat higher than $\Delta G^{\circ}_{VH} \approx 2.8$ Kcal·mol^{-1}

Table 7-3

Apparent Thermodynamics of Defective Helix Formation in 0.1 M NaCl, TE.

Polymer	T_m (K°)	ΔH(Kcal·mol^{-1})	ΔS(cal·mol^{-1})
poly d(GGA)·poly d(TCTC)	320.5	-4.5 ± 20%	-14 ± 20%
poly d(GAA)·poly d(TTT)	300.5	-3.1	-10
poly d(AAA)·poly d(TTC)	299.5	-3.3	-11
poly d(GAA)	304	-0.7	-2.3
poly d(GA)	322.5	-1.9	-5.9
poly d(GGA)	335	-2.0	-6.0

Note: Hybrid data is reported as Kcal·mol bp^{-1} while polypurine data is reported as Kcal·mol purine Pi^{-1} (see text).

reported for extrahelical perthalic acid oxidized A's in a poly r(A,A*)·poly r(U) helix (Fink and Crothers, 1972). By a similar calculation T·G or A·C pairs destabilize ΔH by ≃ 10 Kcal·mol pair^{-1} and $\Delta G°$ by 2.5-2.7 (T·G) and 2.7-3.0(AC) Kcal·mol pair^{-1} depending on reference duplex. The entropy of hybrid formation is less unfavourable than of duplex formation indicating, as one might expect, that these are less ordered structures. The cooperative melting length was circa 130 ± 60 bp.

Further comparisons between the different types of helix defect requires better data.

Measurements of Charge Density

A theoretical relationship exists between residue spacing, transition enthalpy and the dependence of T_m on salt concentration. Is it correct and can one detect the presence of an extrahelical base by the higher charge density it should engender? Table 7-4 summarizes these results calculated on the basis of this theory.

Fibre diffraction data indicates average residue spacings of 1.63 Å (poly dA·poly dT) or 1.60 Å (poly d(TC)·poly d(GA)) depending slightly upon conditions (Leslie et al, 1980) so clearly theory is in considerable error. Assuming b_{coil} is wrong the error can be eliminated if b_{coil} = 2.9-3.1 Å. There is, however, no other experimental data to justify this ad hoc correction. It is still probably justified to intercompare the data in Table 7-4 but even so there is no obvious evidence of a higher charge density in poly d(GGA)·poly d(TCTC) or the tetraplexes. Unfortunately a 20% higher charge density is only going to be barely detectable if b_{helix} = 1.6 ($\Delta b \simeq 0.2$) so although the approach may still be a valid one it will require far more accurate calorimetric data than can be collected at present.

Table 7-4

Calculated Residue Spacings in 0.10 M NaCl, TE.

Polymer	b(Å)	Polymer	b(Å)
poly dA·poly dT	2.8	poly d(GGA)·poly d(TCTC)	3.2
poly d(TTC)·poly d(GAA)	3.2	poly d(GAA)·poly d(TTT)	3.1
poly d(TC)·poly d(GA)	3.0	poly d(AAA)·poly d(TTC)	3.2
poly d(TCC)·poly d(GGA)	3.1		
poly dG·poly dC	3.4	poly d(GAA)	3.5
		poly d(GA)	2.6
		poly d(GGA)	3.3

Note: An average residue spacing of 2.8 Å implies each base-pair is separated by 5.6 Å since there are 2 phosphates per base-pair.

Frameshift Mutagenesis

One final and interesting calculation is:

$$\Delta G_{extrahelical} = \Delta G_{hybrid} - \Delta G_{perfect\ duplex} \quad \text{at } T = 37°$$

$$= -0.42(\pm 0.13) + 3.59(\pm 0.10)$$

$$= 3.2 \pm 0.16$$

so

$$K_{eq} = 6 \pm 0.3 \times 10^{-3} \ .$$

This is a crude, upper thermodynamic limit on the frequency of extrahelical base formation. Coincidentally it is similar to the observed upper limit of spontaneous deletion mutation frequencies observed in the T4 rII locus 2.3×10^{-3} (Pribnow et al, 1981). By comparison thermodynamics does not predict successfully the observed rate of transition mutations. K_{eq} between perfect and defective helicies containing G·T or A·C pairs can be calculated to be $1.6\text{-}2.4 \times 10^{-2}$ depending on the reference helix. This is far too high for known spontaneous transition frequencies of $1\text{-}4 \times 10^{-5}$ in mutD5 E.coli which is presumed to be exonuclease - and hence checking function - deficient (Topal and Fresco, 1976a). The simplest explanation for this is that G·T and A·C pairs stacked inside helices cannot be compared with mismatched pairs being added to the ends.

Drug Binding

Intercalating drugs are characteristic in being able to induce and revert frameshift mutations (Drake and Baltz, 1976). In the Streisinger model this is due to such drugs being able to specifically stabilize configurations containing extrahelical bases. In support of this Lee and Tinoco (1978) observed that ethidium bromide was capable

of stacking into complexes of GpUpG·CpC and CpUpG·CpUpG stabilizing the extrahelical U's. What effect do intercalating drugs have on the stability of a polymer containing extrahelical bases?

Figure 7-6 illustrates the effect of the strongly mutagenic drug proflavine on the melting point of poly d(GGA)·poly d(TCTC). A significant degree of stabilization was observed down to 0.3 μM and the hybrid appeared to be more strongly stabilized than the duplex. This is a rather impressive display of the phenomenon of telestability (Burd et al, 1975) since at 0.3 μM about 1 drug molecule per duplex is still significantly stabilizing the whole structure. The most interesting observation is that above 10 μM proflavine altered the hybrid structure dramatically. The hybrid transition disappeared and a new one appeared melting at lower temperatures. What's remarkable about this is that in vivo dye induced frameshift frequencies are peculiarly concentration dependent. Proflavine concentrations of up to circa 20 μM enhance mutation rates but higher concentrations then cause a decrease in mutation (Drake, 1970). It remains to be proved that the unidentified structure lacks extrahelical bases but if so this provides a rationale for this unexplained phenomenon.

Terbium Fluorescence

Certain lanthanide ions show a striking enhancement of intrinsic fluorescence upon binding to aromatic compounds. Of these lanthanides Tb(III) is perhaps the most widely used fluorescent probe and has been used in both protein and nucleic acid studies. In polynucleotides terbium interacts primarily with guanine residues (Ringer et al, 1978). Binding involves both the base and phosphate since it is residue specific and ionic strength dependent (Yonuschot and Mushrush, 1975).

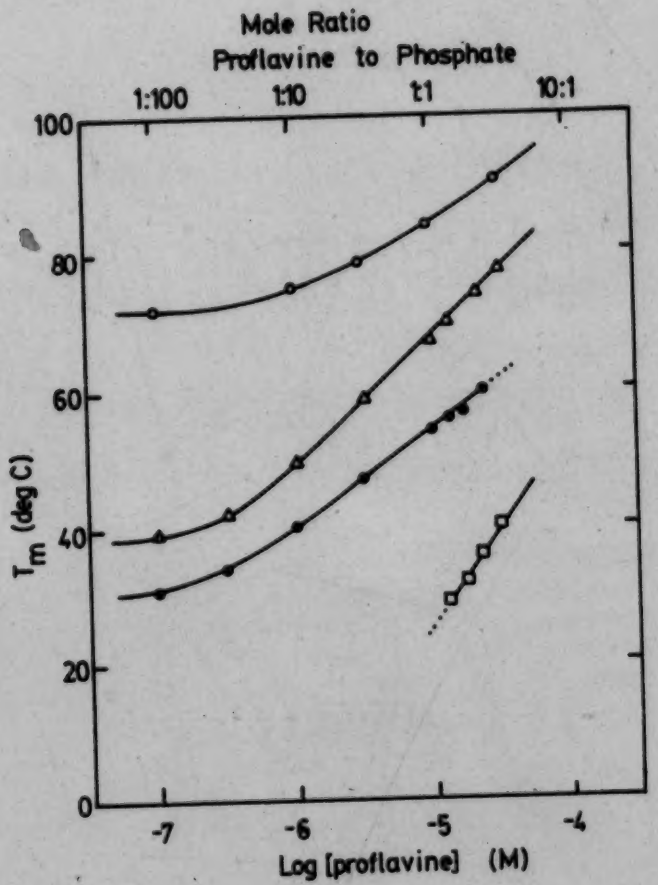


Figure 7-6 Effect of proflavine concentration on hybrid melting. Samples contained 10 μM (Pi) poly d(TCC)·poly d(GGA) or 10 μM (Pi) poly d(GGA)·poly d(TCTC), 5 mM NaCl, TE and proflavine sulphate as indicated in 300 μl. T_m's were recorded spectrophotometricly. (-o-) poly d(TCC)·poly d(GGA), (-Δ-) poly d(GGA), (-●-) poly d(GGA)·poly d(TCTC) and (-□-) unidentified transition. Controls containing only proflavine showed no cooperative melts.

Although poly G gives the strongest fluorescence enhancement other single-stranded polymers will also do so to a lesser extent. Where terbium fluorescence is of particular interest is that base-paired polynucleotides have been reported to give no enhancement of fluorescence even though terbium still binds strongly (Topal and Fresco, 1980). Thus terbium fluorescence methods provide a technique complementary to ethidium bromide fluorescence measuring exactly what ethidium does not.

Controls showed that terbium fluorescence is indeed a sequence dependent phenomenon. Figure 7-7 shows the excitation spectra of free terbium, 1 μM poly dT, 1 μM poly d(GA) and 1 μM poly d(GGA). The sequence dependency is not a simple one since 1 μM poly d(GAA) gives the same fluorescence enhancement as poly d(GA) even though there are only 2/3 the G's present. Sensitivity was adjusted to give the same scale as is used in the ethidium assay thus terbium fluorescence is 1/2 to 1/20 as sensitive at this pH. Optimal settings were found to be excitation 265 nm and emission 545 nm. One feature worth noting in Figure 7-7 is the polynucleotide insensitive excitation peak centred at 225 nm which provides an internal sensitivity calibration.

Terbium fluorescence is ideal for investigating polypurine structure. Poly rG denatured in urea shows half the fluorescence of ordered poly rG (Topal and Fresco, 1980) and this is also true of poly d(GGA). Heat denatured poly d(GGA) added directly to the fluorescence buffer gave twice the fluorescence of poly d(GAA) first annealed in 0.050-0.200 M NaCl. Poly d(GGA) annealed in salt concentrations of 0.7 to 2.0 M also gave half the fluorescence of denatured material. This strongly suggests that although no poly d(GGA) melting transition is observed spectrophotometrically above 0.41 M NaCl the structure is still forming.

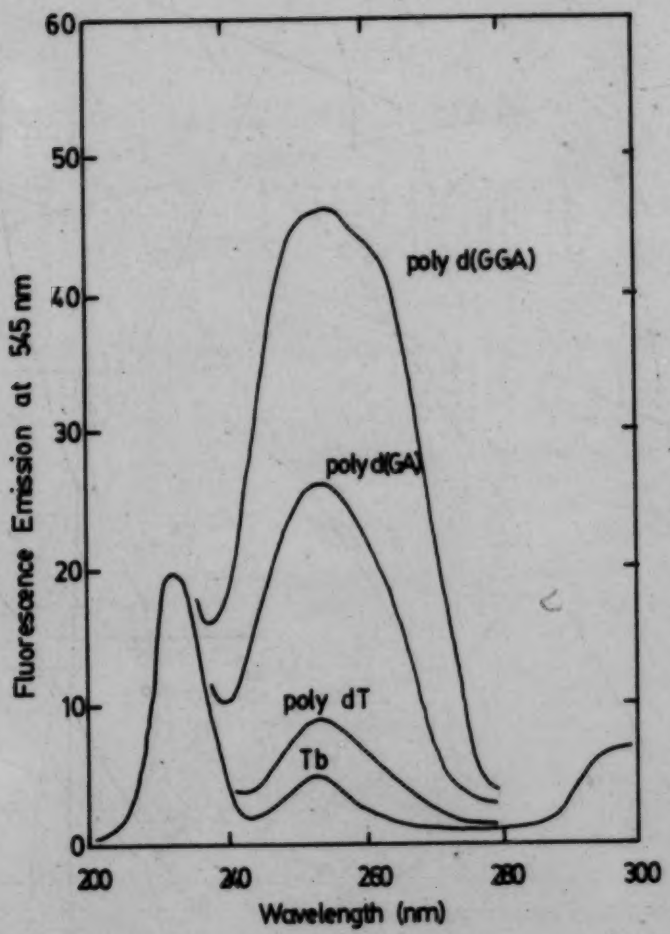


Figure 7-7 Terbium fluorescence spectra. 2ml samples contained 0.5 mM terbium solution with or without 1 μM indicated polynucleotide. Blank was distilled water in all cases. The "peak" at 360 nm is a doubling mirror artifact.

The observation that poly d(GGA) behaves as does poly rG is further evidence of similar structures.

Terbium fluorescence can also be used to measure mixing ratios since formation of a duplex results in strong fluorescence quenching. Formation of poly d(GGA)·poly d(TCC) from the tetraplex and poly d(TCC) was accompanied by an 85% drop in fluorescence and mixing ratios agreed perfectly with ethidium derived ratios. Unfortunately the drop in fluorescence accompanying hybrid formation was not nearly as dramatic, 10% for poly d(GAA)·poly d(TTT) and 20% for poly d(GGA)·poly d(TTC) for example. This made estimating endpoints very difficult and the errors large. This is probably due to the low stability of these structures since a mismatched pyrimidine strand cannot be expected to compete as effectively as can the correct strand with 0.5 mM terbium for guanine binding site(s). pH is also a problem since protonation of C's starts to occur to a significant extent below pH 7 (depending on salt). A pH of 6.3 is a compromise between sensitivity [maximal at pH 5.5 (Yonuschat and Mushrush, 1975)], helix protonation (pH preferably 7 or greater) and terbium stability (hydroxides precipitate at pH's of 7 or greater). However with better instrumentation, lower terbium concentrations and higher pH's this technique shows promise of being a useful one for probing hybrid structure.

Conclusions

A variety of physical techniques have been used to investigate the properties of hybrid polynucleotides containing mismatched and extrahelical bases. Hybrids containing large numbers of helix defects are sufficiently stable that hybridization can be demonstrated by buoyant density studies. This suggests that random copolymers can alter the buoyant densities of denatured DNA's by annealing with sequences that

aren't strictly complementary. NMR shows that at least two distinct phosphodiester conformations exist in poly d(GGA)·poly d(TCTC) but dramatically shifted extrahelical base phosphate resonances aren't observed. Calorimetry provides quantitative information concerning the instability of these hybrids. Extrahelical bases appear to be marginally more destabilizing than are mismatched bases which agrees with measured destabilization indices. From thermodynamic data it is possible to derive upper limits on the predicted frequency of frameshift mutations that agree with observed frequencies but one cannot predict observed missense and nonsense mutation rates. This indicates that hybrids containing extrahelical bases are suitable substrates for investigating the origins of frameshift mutations but not transition or transversion mutations. This being so the effect of proflavine on poly d(GGA)·poly d(TCTC) was examined and the unusual effects it has on hybrid stability may rationalize a number of unexplained features of drug induced mutagenesis. Finally terbium fluorescence appears to be a promising technique for further investigations into hybrid structure.

VIII. Biological Studies

Hybrids containing unusual structures are potentially useful substates for investigating enzyme specificity and protein function. Chapter VII summarizes some preliminary studies.

Human Placental DNA Binding Proteins

Over the past 15-20 years a wide variety of enzymes have been identified and purified whose specific function is to repair DNA damage. Such enzymes have been isolated from a variety of organisms ranging in complexity from bacteriophage to man. In man a protein has been isolated from placental tissue which is distinguished functionally by a specific affinity for ultraviolet-damaged DNA (Feldberg and Grossman, 1976; Feldberg, 1980; Feldberg (in press)). It has no observed enzymatic activity and the substrate-specificity is unknown. The lesions recognized are not pyrimidine dimers. The wide variety of agents in addition to ultraviolet that produce binding sites suggest that this protein recognizes a structural alteration common to many types of DNA damage. A helix distortion such as that associated with mismatched bases is a possible candidate (Feldberg, 1980). If this hypothesis is true hybrid polynucleotides should also bind this protein.

The human placental binding protein is normally assayed by the ability to retain ultraviolet treated radioactive T7 DNA on nitrocellulose filters (Figure 8-1). DNA-protein complexes stick while free DNA passes through. By this criteria the samples provided by Dr. Feldberg survived freezing and thawing with circa 80% recovery of activity. This assay did not seem suitable for assaying binding to hybrids since hybrid polynucleotides are likely to contain single-stranded

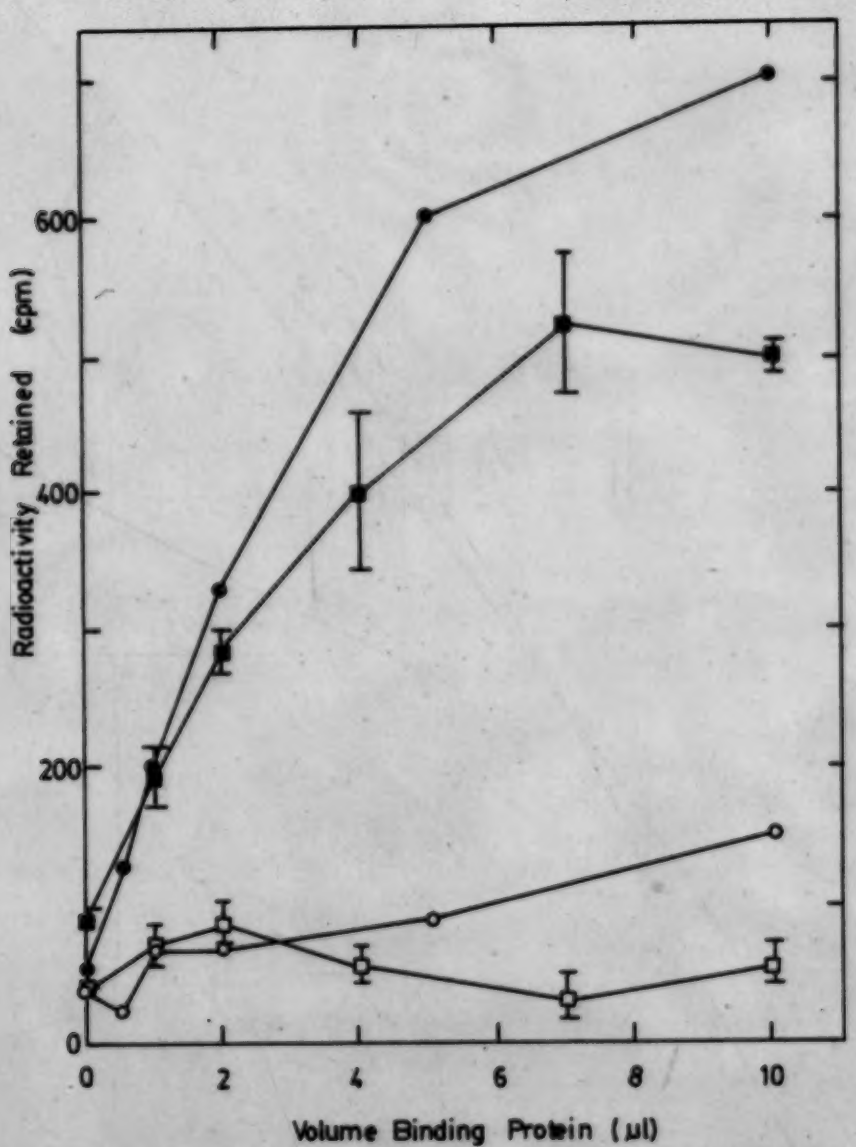


Figure 8-1 Human placental DNA binding protein assay. 300 µl reactions contained 10 mM Tris·HCl pH 8, 50 mM NaCl, 1080 cpm [^{3}H] labelled T7 DNA (0.9 µM) and the indicated volume of "Ha-ultrogel" or "Seph IV" fraction binding protein. After incubation (1 hour at 22°), 2 ml ice-cold 2 x SSC (0.3 M NaCl, 30 mM Na·citrate) was added, rapidly filtered through nitrocellulose and filters washed with a second 2 ml volume of 2 x SSC. After drying samples were counted. UV irradiated T7 DNA was prepared by exposing under a 254 nm lamp for 2 min. (200 J/m^2, ~13 sites per molecule; Feldberg, 1980). Seph IV fraction (-●-) irradiated DNA, (-o-) control DNA; HA-Ultrogel fraction (-■-) irradiated DNA; (-□-) control DNA.

regions and single-stranded DNA's also stick to nitrocellulose. Because of this problem I first attempted to develop a solid-phase binding assay using DNA adsorbed onto a solid support and ^{125}I labelled binding protein. This seemed a reasonable approach since Dr. Lee was using a similar assay in his experiments on monoclonal antibody production (Lee et al, 1981). Unfortunately, when no binding specificity was observed under a variety of conditions these experiments were abandoned.

Nitrocellulose assays will work with synthetic DNA's if one pretreats the filters with alkali (eg. McEntee et al, 1980). This procedure prevents single-strand binding yet DNA·protein complexes are still retained. However when no binding to hybrids such as poly d(GGA)·poly d(TTC), poly d(GAGAGA)·poly d(TCCTCC) or bisulfite treated poly d(TCC)·poly d(GGA) was observed it became apparent that this assay is not sensitive enought to measure binding to very short polymers. It is reasonable to assume that one bound protein is capable of trapping an entire molecule of DNA (Feldberg, 1980) which is about 40,000 base-pairs if T7 DNA is used, but only about 50 base-pairs if synthetic DNA's are used. This meant that ^{3}H labelled synthetic polymers had only about 1/2000 the specific activity of T7 DNA. Using a ^{32}P end-label it was possible to bring the specific activity of synthetic polymers up to about 1/200 of the ^{3}H labelled T7 but this was still insufficient to detect binding. These experiments need longer polymers or polymers synthesized with [α] ^{32}P dXTP's before definite conclusions can be drawn concerning the substrate specificity of this protein.

Endonuclease Specificity

A few single-strand specific endonucleases were examined to see whether they displayed any ability to discriminate between true

duplex polynucleotides and various types of hybrids. In all these studies, conditions were modified so as to maximally stabilize the hybrids (0.5 M NaCl) while minimizing the possibility of alternative structures forming. Thus the enzyme properties described here are not necessarily the same as those observed under optimal conditions.

Two single-strand specific endonucleases have been isolated from N. crassa conidia (Rabin and Fraser, 1970; Linn and Lehman 1965a,b). Based on the observed elution position on phosphocellulose the enzyme studied here appears to be the smaller (55000 M_r) one of Linn and Lehman. Of relevance to these experiments is the 50-100 fold preference for denatured polynucleotides and its sequence specificity (digestion products appear in the order dGMP > dAMP, dTMP > dCMP). Figure 8-2 shows the digestion of poly d(GA) hybrids. No obvious correlation between nuclease resistance and T_m or presumed structure was noted in these or other experiments. What is shown clearly is the sequence specificity of this enzyme. Free polypyrimidines are digested to acid soluble products in the order poly d(TTC) > poly d(TC) > poly d(TCC) which is as expected considering earlier studies on this enzyme (Linn and Lehman, 1965b). This sequence specificity also seems to be reflected in the rate of hybrid degradation (Figure 8-2) but a direct comparison of single-strand and duplex digestion rates is not possible. Such specificity makes the N·crassa endonuclease unsuitable for probing polynucleotide structure.

BAL endonuclease is isolated from supernatents of Pseudomonas BAL 31 cultures (Gray et al, 1975). It is primarily a single-strand specific endonuclease although it will slowly exonucleolytically degrade duplex polymers. The endonuclease appears able to detect helix distor-

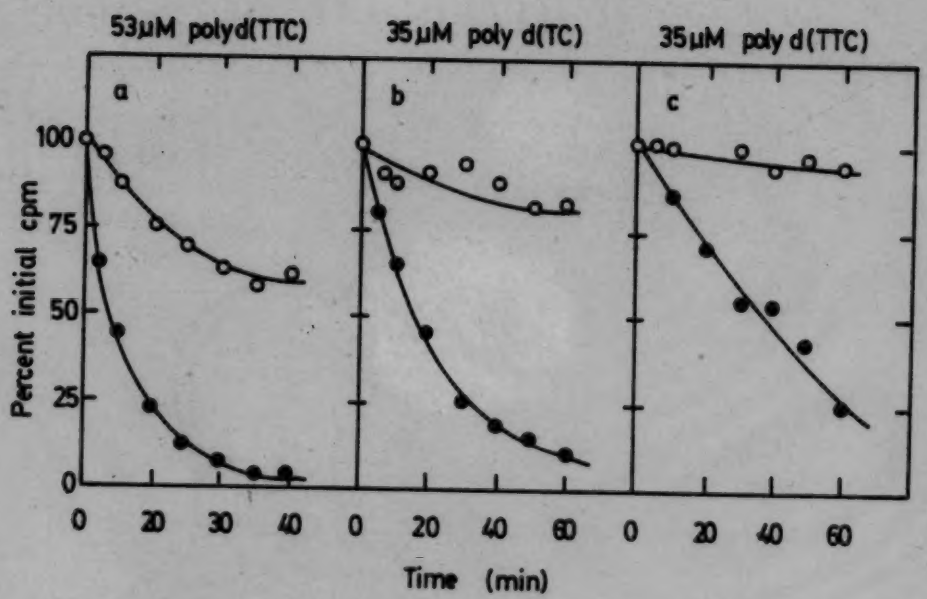


Figure 8-2 N.crassa endonuclease digestion of poly d(GA) hybrids. Reaction (100 µl) contained 0.5 M NaCl, 0.1 M Tris pH 8.0, 10 mM $MgCl_2$, 2.7 U N.crassa endonuclease and ^3H labelled polypyrimidine either with (-o-) or without (-•-) 35 µM poly d(GA). 100% was 2400-2800 cpm. 10 µl samples were TCA precipitated and counted.

Reactions were saturated with substrate since doubling its concentration had no effect on the rate of single-strand digestion. Note that EDTA inhibits N.crassa endonuclease at concentrations well below those needed to stoichiometrically chelate Mg^{+2} (Linn and Lehman, 1965a). EDTA concentrations were 71, 61 and 58 µM in a, b and c, respectively, under which conditions the enzyme is 90-95% inhibited.

tions since it will nick supercoiled but not relaxed covalently closed circular DNA's. This is presumably due to enzyme attack or transiently unwound or cruciform structures. The enzyme also nicks relaxed circular DNA's modified with N-acetoxy-N-2-acetylaminofluorene or nitrous acid. In both cases not all sites appear to be substrates. For example nitrous acid causes both cross-linking and deamination but probably only the former is a substrate (Legerski et al, 1977). These observations suggest that the enzyme is capable of recognizing gross disruptions in helix geometry and should be able to distinguish hybrids from proper duplexes.

BAL 31 endonuclease did show a considerable capacity to discriminate between hybrid polynucleotides and true duplexes. Figure 8-3 shows the digestion of poly d(GA) and poly d(GGA) containing hybrids. Proper duplexes were fairly resistant while hybrids were rapidly degraded. Again no obvious correlation between sensitivity and T_m or presumed structure was noted. At present it is difficult to say whether this sensitivity is simply due to helix distortions being attacked per se or is due to nuclease digestion of transiently single-stranded regions formed by breathing and strand-migration. An examination of the digestion products might resolve this ambiguity. These results suggest that further experiments may be worth pursuing.

S_1 endonuclease (Vogt, 1973) was also tested and, for example, poly d(GGA)·poly d(TCTC) was observed to be quite resistant compared to single-stranded poly d(TC). The interpretation of such experiments is difficult since they were performed at pH 7.0 and the S_1 pH optimum is 4.2.

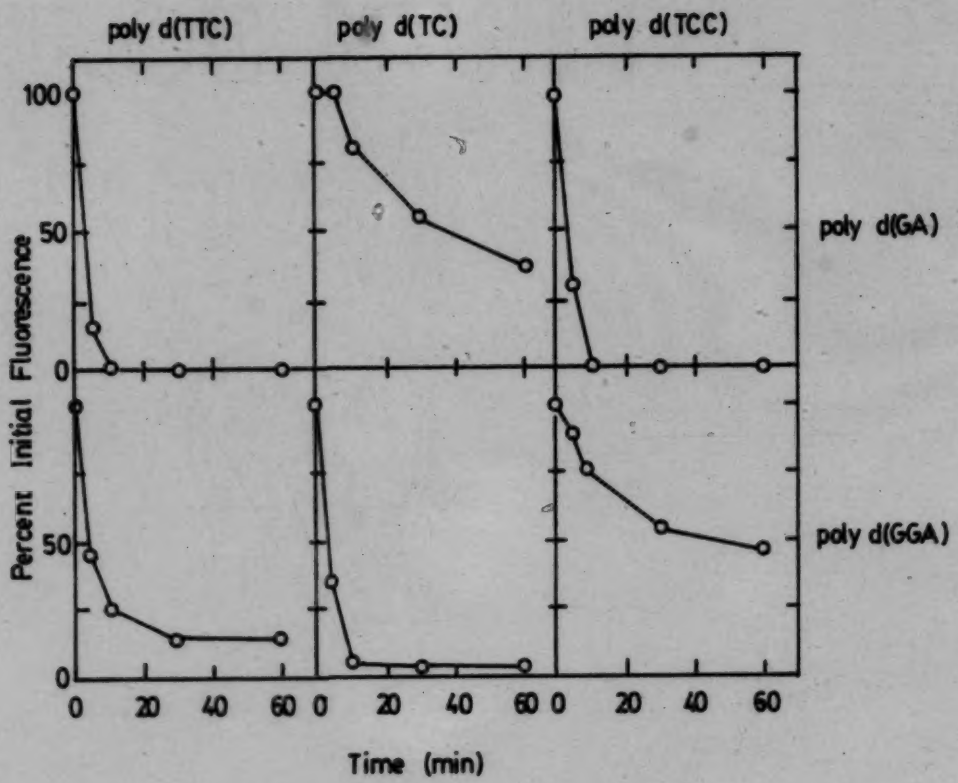


Figure 8-3 BAL 31 endonuclease digestion of poly d(GA) and poly d(GGA) hybrids. Reactions (55 µl) contained 70 µM DNA (in stoichiometries as described in Chapters V and VI), 0.5 M NaCl 0.1 M Tris pH 8.0, 1 mM $MgSO_2$, 1 mM $CaCl_2$ and 0.41 µg BAL endonuclease. 10 µl samples were added to 2.0 ml pH 8 ethidium solution and fluorescence measured. 100% was 19-29 fluorescence units depending on hybrid.

RNA Polymerase

Although under normal conditions RNA polymerase is known to transcribe DNA with considerable accuracy (Springate and Loeb, 1975) under special conditions (particularly in the presence of Mn^{+2}) it catalyses some quite unusual reactions. These include the de novo synthesis of poly rA·poly rU from ATP and UTP in the absence of template (Smith et al, 1967) and synthesis of homopolymers off non-complementary single or double stranded templates (Nishimura et al, 1965; Adman and Grossman, 1967; Paetkau et al, 1972).

RNA polymerase will also catalyse the synthesis of poly rA off poly d(TTC) and poly rC off poly d(GGA) (Figure 8-4). A poly rC + poly d(GGA) hybrid could be prepared but a poly rA + poly d(TTC) hybrid could not. These and other results suggest that while the ability of RNA polymerase to catalyse synthesis of non-complementary polymers is related to the ability to form mismatched hybrids (Paetkau et al, 1972) other factors such as hybrid stability must also be important. Considering the complexity of RNA polymerase catalysed reactions (Losick and Chamberlin, 1976) it is going to take more sophisticated experiments than these to exactly define the role of non-Watson-Crick base-pairing in this peculiar phenomenon.

Conclusions

Some potential exists for using hybrid polynucleotides for investigating enzyme function but clearly a problem associated with these experiments is that one has to better define substrate structure before one can make conclusions regarding enzyme function. Thus the observed sensitivity to BAL endonuclease may be rationalized in two different ways because of uncertainties regarding hybrid stability.

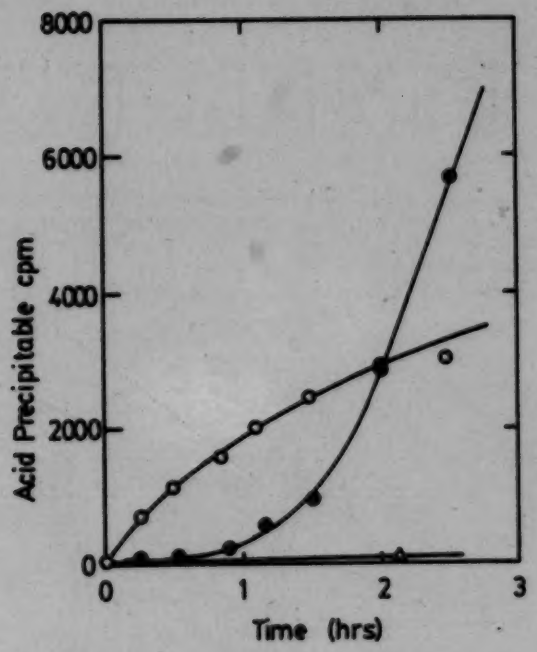


Figure 8-4 Synthesis of homopolymer RNA's off non-complementary single-stranded DNA's by RNA polymerase. Reactions (100 µl) contained 40 mM Tris·HCl pH 8, 1.6 mM $MnCl_2$, 4.8 mM Mg Cl_2, 10 mM β-Mercaptoethanol, 0.2 mM XTP ([^{14}c] labelled 0.4 m Ci/m mol), 20 µM DNA and 0.2 mg/ml RNA polymerase (Paetkau et al, 1972). 10 µl aliquots were TCA precipitated at the indicated times and counted. (-o-) poly rA off poly d(TTC); (-●-) poly rC off poly d(GGA); (-Δ-) poly rC off poly d(GA). Blank (no template) was identical to poly d(GA) primed reaction. Note the difference in lag times. Poly rC + poly d(GGA) hybrids form but poly rA + poly (TTC) and poly rC + poly d(GA) hybrids do not as assayed by fluorescence.

The same is true of the RNA polymerase data. Furthermore one must clearly be aware of potential artifacts associated with the particular physical properties of these polymers. Failure to bind the human placental binding protein was due to polymer size not polymer structure and even had binding been observed it would remain to be shown that some minor product of strand preparation wasn't the actual substrate. Further experiments are clearly worth pursuing but the interpretation is going to be less straightforward than would appear at first glance.

IX. Conclusions and Further Experiments

It is apparent that conformation is a highly flexible parameter and in order to maximize base-pairing and base-stacking a variety of defects can be accommodated into base-paired helices. These defects include extrahelical bases, G·T pairs and probably also A·C pairs. Many of the structures formed are surprisingly stable and can in all likelyhood exist under physiological conditions. Because they are stable this suggests that these hybrids may be relevant to our understanding of biological processes.

Without repeating in detail the conclusions presented at the end of each preceding chapter the following points seem worth recalling.

DNA Polymerase I

DNA polymerase I is sensitive to the presence of DNA binding drugs and this sensitivity can be used to advantage to inhibit the synthesis of poly d(AT). It is still not clear what is so unusual about poly d(AT) that it is synthesized in preference to all other polymers or why in particular poly d(TC)·poly d(GA) should be synthesized de novo under special conditions when other polymers are not Furthermore it is not clear how the 5'-exonuclease is involved in this phenomenon. The Klenow polymerase fragment (which lacks the 5'-exonuclease) reportedly does not catalyse de novo poly d(AT) synthesis (Setlow, Brutlag and Kornberg, 1972).

Drug Binding Specificity

The ability of TANDEM to bind poly d(TAC)·poly d(GTA) indicates that this duplex contains the same binding site found in poly d(AT). This suggests that the sequence recognized is TpA. This may

indicate that poly d(TAC)·poly d(GTA) contains the dinucleotide repeat unit that is present in poly d(AT) (Klug et al, 1979). If this could be shown (by ^{31}P NMR) it would be an interesting demonstration that DNA's can contain local substructures determined by local base sequences. Poly d(ATC)·poly d(GAT) seems to lack this site demonstrating how base orientation can influence ligand binding.

Proflavine sulphate shows an intriguing ability to stabilize and destabilize hybrid polynucleotides containing extrahelical bases but it remains to be shown that the structure formed at high proflavine concentrations lacks extrahelical bases. Intercalating drugs differ in their mutagenicity (Drake, 1970) and it would be interesting to see if there is any correlation between mutagenicity and ability to stabilize extrahelical bases. Scatchard analysis might provide some clues regarding how these compounds stabilize extrahelical bases since several mechanisms seem possible (Lee and Tinoco, 1978).

Hybrids Containing Extrahelical Bases

A variety of hybrids can be formed containing extrahelical bases. The best characterized is poly d(GGA)·poly d(TCTC) which contains extrahelical T's. In DNA·RNA hybrids extrahelical bases can also occur but seem to appear only on the RNA strand. These observations cannot at present be generalized since only a small number of hybrids have been studied in detail. Other combinations of strands remain to be examined. These include a variety of non-pyrimidine·purine DNA's, DNA·RNA hybrids and RNA·RNA hybrids. Studies on RNA·RNA hybrids would be of considerable interest since accurate thermodynamic data would make possible better RNA folding algorithms (Salser, 1977). These use van't Hoff derived parameters at present some of which may

be in error by ±100% (Tinoco et al, 1973). Studies on other hybrids should make possible generalizations concerning the effect of surrounding sequence on defect stability.

Hybrids containing extrahelical bases show physical properties that are consistent with observation on frameshift mutagenesis. Calorimetry shows that they destabilize helices enthalpically but stabilize entropically. Extrahelical bases appear to destabilize duplexes slightly more than do base mismatches.

Hybrids Containing Base Mismatches

Hybrids containing G·T, A·C and pyrimidine·pyrimidine oppositions can also be prepared. While it seems likely that G·T and perhaps A·C pairs occur, it is difficult to imagine how T·T or C·T pairs can form between G·C pairs. These defects also destabilize enthalpically and stabilize entropically but do not appear to be good models for transition and transversion mutation.

A subject which has not been discussed is the importance of keto-enol and amino-imino tautomerization. Do mismatched structures contain wobble or tautomeric base-pairs? It is a widely held belief that some mutations originate with tautomeric non-Watson-Crick pairs (e.g. Topal and Fresco, 1976a) yet what little physicochemical data there is on the subject is not only misleading but wrong (Drake and Baltz, 1976). Wolfenden (1969) is still widely quoted regarding tautomeric equilibria in inosine and adenine yet the experimental and theoretical assumptions implicit in that paper cannot be justified. Studies on bases in solution are going to be of dubious value anyway when extrapolated to stacked helices or polymerase active sites.

These hybrids provide an opportunity for investigating tautomeric frequencies in polymers but methods other than UV spectroscopy will have to be used. Because aromatic hydrogen-bonded protons appear well down-field of most resonances and show unique exchange properties (Early et al, 1978) proton NMR appears to be the method of choice for these studies.

The poly d(GGA)·poly d(CCC) hybrid was curious because above a certain salt concentration the optical melting transition disappeared. Calorimetry might show that the hybrid was still melting - a straightforward demonstration that not all polymer transitions are accompanied by optical transitions.

Duplex Polynucleotide Structure

Calorimetry shows that the stability of polypyrimidine·polypurine DNA's is dependent on base composition. Because of experimental scatter it is not clear whether this stability can be predicted simply on the basis of percent G·C composition or whether nearest neighbour effects are also involved. Based on T_m measurements polypyrimidine·polypurine DNA's are known to be less stable than other DNA's (Wells et al, 1970) which may be because interhelical stacking cannot occur. More extensive calorimetric measurements using pyrimidine·purine DNA's and other DNA's may make it possible to separate out these effects.

As mentioned earlier the structure of poly d(TAC)·poly d(GTA) also seems worthy of further investigation.

Biological Aspects

Chapter VIII indicates how these hybrids might be used in investigating DNA binding protein and nuclease specificity. It is

also apparent that RNA polymerase catalysed non-complementary RNA synthesis is related to hybrid formation but is a more complex phenomenon than might first appear.

An advantage of these hybrids is that high frequencies of defined helix defects can be otained. Dr. Lee has used poly d(GAA)·poly d(TCTC) hybrids as antigens and observed significantly stronger responses in mice to this hybrid than to ordinary duplexes lacking extrahelical bases. These might make possible a rapid method for selecting genetically defined lesions in heteroduplex restriction fragments if antibodies with a sufficiently high binding constant can be found. Obviously one is not restricted to extrahelical bases and these hybrids may be suitable immunogens for generating antibodies to a whole spectrum of DNA defects.

Whether these hybrids are of any relevance in other biological systems remains unknown. Leaky translational frameshifting (Atkins et al, 1979), spontaneous frameshifting during replication (Streisinger et al, 1966), heteroduplex formation during recombination (Das Gupta and Radding, 1982) and base damage (e.g. Karran and Lindahl, 1980) all may involve or generate extrahelical and mismatched bases. Whether these hybrids might be of use in studies concerning repair and recombination is a question that further work must resolve.

Appendix I.

Monovalent Cation Effects on Polynucleotide Stability

Nucleic acids are highly charged polyelectrolytes and a major destabilizing force will be due to electrostatic repulsion between adjacent phosphates. This force tends to denature double and multistranded polynucleotides through interstrand repulsion and modify their overall geometry by intrastrand repulsion. The importance of counterions is clear; removal of them progressively lowers the melting point of unprotonated multistranded polynucleotides to the point that in "counterion free" water they will spontaneously denature.

The most successful theory describing monovalent cation dependent effects is that of Manning (Manning, 1972a,b; 1978) as modified and extended by Record et al (1976, 1978). Since polynucleotides are negatively charged polyanions cations will associate so as to neutralize most of the phosphate charges. The extent of association is believed to be purely a function of charge density (Manning, 1972a). Counterions associate in two distinct ways by Debye-Hückel shielding and by condensation of counterion directly onto the DNA. In the limiting case of low counterion concentration (<1M for duplex DNA's, <0.1 M for single stranded DNA's) the amount of charge neutralized in the second way can be calculated by

$$\psi_c = 1 - \xi^{-1} \qquad \text{(i)}$$

where ψ_c is defined as the fraction of phosphate charged neutralized by counterion condensation and ξ is defined by

$$\xi = \frac{e^2}{n\epsilon kTb} = \frac{7.14}{nb}\ \text{Å} \quad \text{(in } H_2O \text{ at } 25°\text{)} \qquad \text{(ii)}$$

Here e = electron charge, n = cation valence, ε = bulk dielectric constant, T = absolute temperature [The product εT turns out to be nearly independent of temperature over the range of temperatures used here (Manning, 1972a,b] and b = average charge spacing as measured along the helix axis. A further fraction of the charge will be neutralized by a diffuse cation cloud. This can be calculated by (Record et al, 1976)

$$\psi_s = (2\xi)^{-1} \qquad \text{(iii)}$$

where ψ_s is now defined as the fraction of charge neutralized by Debye-Hückel screening. Combined these two equations give the total charge neutralized

$$\psi = \psi_c + \psi_s$$
$$= 1 - \xi^{-1} + \frac{\xi^{-1}}{2} = 1 - \frac{\xi^{-1}}{2} \qquad \text{(iv)}$$

For double stranded DNA in solution b is 1.7 Å (Record et al, 1978) hence we can calculate

$$\psi = \psi_c + \psi_s$$
$$= 0.76 + 0.12$$
$$= 0.88$$

Thus the amount of net charge remaining per phosphate has been reduced by these two effects from -1.0 to -0.12. The importance of this can be easily (albeit crudely) calculated. The work required to bring two charges $q_o = e$ to within 1.7 Å of each other can be calculated by (Metzler, 1976)

$$W = 2.16 \times 10^7 \times \frac{q_o^2}{r\varepsilon}$$
$$= 332\ (\varepsilon r)^{-1}\ \text{(Kcal/mol)} \qquad \text{(v)}$$

where r = separation (Å)

ε = solvent dielectric

this turns out to be about 2.4 Kcal/mol paired charges. After counterion condensation these have been reduced to 0.24 q_o hence $W = W_o(0.24)^2$ = 0.14 Kcal/mol a considerable reduction in energy. Screening will further reduce this.

Theory such as this predicts a number of experimental results. Counterion condensation should not depend upon bulk ionic strength since the amount of counterion condensed depends only upon charge density. Furthermore in a thermodynamic sense charges are charges and consequently the extent of counterion binding should be independent of counterion identity although dependent on valence. Assuming ^{23}Na NMR to measure only condensed counterions Anders et al (1978) measured $\psi_c = 0.75 \pm 0.1$ [the experimental uncertainty is probably higher see Bleam et al (1980)] and clearly independent of salt concentration over the range 0.003 - 1.3 M (Bleam et al, 1980). Furthermore the amount of salt bound appears independent of base composition (Shapiro et al, 1969) and monovalent cation type although the relative affinities of DNA for cations does seem to vary slightly. With the exception of this last observation the results support a simple electrostatic model of monovalent cation binding to DNA.

DNA Melting

Increasing the counterion concentration stabilizes DNA against thermal denaturation. Why? Consider the fractional charge remaining on a denatured ("coil") polynucleotide in which b shall be taken to be 4.3 Å (experimentally determined as discussed below). Then

$$\psi_{coil} = \psi_c + \psi_s$$
$$= 0.40 + 0.30$$
$$= 0.70$$

and the transition from helix ($\psi = 0.88$) to coil ($\psi = 0.70$) is associated with the release of

$$\Delta\psi = \psi_{helix} - \psi_{coil}$$
$$= 0.88 - 0.70$$
$$= 0.18$$

0.18 counterions per phosphate. By Le Chatelier's principle increasing the salt concentration will stabilize the form associated with more salt.

Experimentally polynucleotide melting enthalpies are observed to be essentially independent of salt concentration (Privalov et al, 1969; Filimonov and Privalov, 1978). Thus salts affect the stability of polynucleotides by altering the entropy of helix formation. Intuitively the stability of DNA duplexes will represent a balance between an unfavourable entropy associated with condensed counterions, a favourable free energy gain associated with the charge neutralization, a favourable enthalpy of base-pairing and base-stacking and unfavourable entropy of helix formation. DNA melts above a certain temperature since higher temperatures selectively magnify the unfavourable entropy terms in

$$\Delta G = \Delta H - T\Delta S \qquad \text{(vi)}$$

Note that the stabilizing effect of salts cannot increase indefinitely. Above a certain salt concentration [typically ~1.2 M (Manning, 1978)] the bulk salt concentration begins to equal that of the DNA associated salt. Consequently T_m is expected to become salt independent at some

point and this is also observed (eg. Wells et al, 1970).

Measurement of Charge Density

The observation that melting of uncharged polynucleotides is associated with counterion release (Ascoli et al, 1961) indicates it should be possible to determine charge densities from thermodynamic data. Consider the free energy charge associated with melting in the following process.

$$DNA_{helix} \rightarrow DNA_{coil} + nM$$

where n is the number coutnerion released per phosphate

$$n = \Delta\psi_c = \xi^{-1}_{helix} - \xi^{-1}_{coil} \qquad \text{(vii)}$$

Expressing DNA concentration as phosphate

$$K_{eq} = \frac{[DNA_{coil}] - [M]^n}{[DNA_{helix}]}$$

$$= [M]^n$$

and

$$\Delta G_s = -RT \ln keg$$

$$= -n\, RT_m \ln M$$

$$= -RT_m \Delta\psi_c \ln M \qquad \text{(viii)}$$

ΔG_s is a purely entropic component associated with release of condensed counterions.

The electrostatic free energy of two charges separated by distance b can be calculated from (Cantor and Schimmel, 1980)

$$G_{el} = -RT\, \xi^{-1} \ln \kappa b \qquad \kappa b << 1 \qquad \text{(ix)}$$

where ξ^{-1} represents the fraction of charge remaining after condensation and κb Debye-Hückel screening parameters. Note that $1 - \xi^{-1} = \psi_c$

and $\kappa = 0.33\ M^{1/2}$ (both are relatively independent of temperature) so

$$G_{el} = -RT\ \xi^{-1}\ \ell n(0.33\ M^{1/2} b)$$

$$= -RT\ \xi^{-1}\ (\tfrac{1}{2}\ \ell n\ M + \ell n\ 0.33b)$$

then ΔG_{el} associated with melting can be shown to be

$$\Delta G_{el} = -\frac{RT_m}{2}\ (\xi^{-1}_{coil} - \xi^{-1}_{helix})\ \ell n\ M - RT_m(\xi^{-1}_{coil} - \xi^{-1}_{helix})$$

$$x\ \ell n\ (\frac{b_{coil}}{b_{helix}})$$

$$= -\frac{RT_m}{2}\ (-\Delta\psi_c)\ell n\ M - RT_m(-\Delta\psi_c)\ell n\ (\frac{b_{coil}}{b_{helix}})$$

$$= \frac{RT_m}{2}\ \Delta\psi_c\ \ell n\ M + RT_m\ \Delta\psi_c\ \ell n\ (\frac{b_{coil}}{b_{helix}}) \qquad (x)$$

Remembering that b is the average charge spacing measured along the helix axis. Now the total free energy change associated with melting will be

$$\Delta G = \Delta G_{el} + \Delta G_s + \Delta H + T_m \Delta S \qquad (xi)$$

where ΔH is the enthalpy change and ΔS the entropy change accompanying melting of each residue. Then summing these terms

$$\Delta G = \frac{RT_m}{2}\ \Delta\psi_c\ \ell n\ M + RT_m\ \Delta\psi_c\ \ell n(\frac{b_{coil}}{b_{helix}}) - RT_m\ \Delta\psi_c\ \ell n\ M + \Delta H + T_m \Delta S$$

$$= -\frac{RT_m}{2}\ \Delta\psi_c\ \ell n\ M + RT_m\ \Delta\psi_c\ \ell n(\frac{b_{coil}}{b_{helix}}) + \Delta H + T_m \Delta S$$

$$= 0 \text{ (at the melting point)}$$

so rearranging

$$-\Delta H = \frac{RT_m}{2}\ \Delta\psi_c \ell n\ M + RT_m\ \Delta\psi_c \ell n(\frac{b_{coil}}{b_{helix}}) + T_m \Delta S$$

$$\frac{\Delta H}{RT_m} = \Delta\psi_c\ \frac{\ell n\ M}{2} - \Delta\psi_c\ \ell n(\frac{b_{coil}}{b_{helix}}) - \frac{\Delta S}{R}$$

or $$\frac{1}{T_m} = \frac{R\Delta\psi_c}{2\Delta H} \ln M - \frac{R\Delta\psi_c}{\Delta H} \ln\left(\frac{b_{coil}}{b_{helix}}\right) - \frac{\Delta S}{\Delta H} \qquad \text{(xii)}$$

Differentiating with respect to salt concentration (and assuming ΔH and ΔS are independent of salt concentration)

$$-\frac{1}{T_m^2} = \frac{R\Delta\psi_c}{2\Delta H} \frac{d}{d} \frac{\ln M}{T_m}$$

or $$\frac{dT_m}{d\ln M} = -\frac{RT_m^2}{\Delta H} \times \frac{\Delta\psi_c}{2} \qquad \text{(xiii)}$$

Which is essentially the same equation as has been derived elsewhere using more sophisticated methods (eg, Manning, 1978; Record et al, 1978; Privalov et al, 1969; Cantor and Schimmel, 1980).

For T4 DNA melting $b_{helix} = 1.7$ Å, $d\,T_m/d\ln M = 8.9°$, $RT_m^2/\Delta H = 50°$ so $\Delta\psi_c/2 = -0.18$ (Record et al, 1978). Since

$$\Delta\psi_c = \xi_{helix}^{-1} - \xi_{coil}^{-1}$$

$$= \frac{b_{helix} - b_{coil}}{7.14}$$

b_{coil} can be calculated to be -4.3 ± 0.2 Å. Conversely it may be reasonable to assume that most "coil" form polynucleotides have similar charge densities so from this b_{helix} could be calculated. Making this assumption its apparent that

$$\Delta\psi_c = -\frac{2\Delta H}{RT_m^2} \times \frac{dT_m}{d\ln M} = \frac{b_{helix} - b_{coil}}{7.14}$$

$$b_{helix} = b_{coil} - 3.12 \frac{\Delta H}{T_m^2} \times \frac{dT_m}{d\log M} \quad (\text{Å}) \qquad \text{(xiv)}$$

Because of limitations in theory (xiv) is only appropriate in salt concentrations less than 0.2M.

Appendix II.

Assumptions in Calorimetry

A. In general the products of melting will not have the same partial specific heat capacities as the reactants. This will result in a break in a plot of C_p as a function of T at T_m. Since scanning calorimetry measures C_p as a function of T this generates a discontinuity in the base-line at $T = T_m$. This is illustrated schematically (Figure A-2(1)) where two linear heat capacity functions are plotted over which is superimposed a melting transition occuring at different temperatures. Clearly melting enthalpies can only be directly compared at some standard temperature in order to account for this difference in partial specific heat capacity. Measurements are normally extrapolated from enthalpies determined at higher temperatures to standard conditions of 298°K. It is assumed throughout the discussion in Chapter VII that there are no partial specific heat capacity differences between duplex and single-stranded DNA's and that enthalpies can therefore be compared at any temperature. This is certainly true within the limits of experimental scatter.

B. In general some single-strand stacking remains after melting of duplex DNA's. Filimonov and Privalov (1978) correct poly rA·poly rU melting enthalpies for residual poly rA stacking. This is necessary in order to derive true enthalpies of helix formation from single strands. Schematically this correction is illustrated in Figure A-2(2). A cooperative duplex melting curve is superimposed on a non cooperative single-strand melt. The true enthalpy of duplex formation from unstacked single strands is calculated by summing the enthalpy of duplex formation plus the hatched curve area. In general

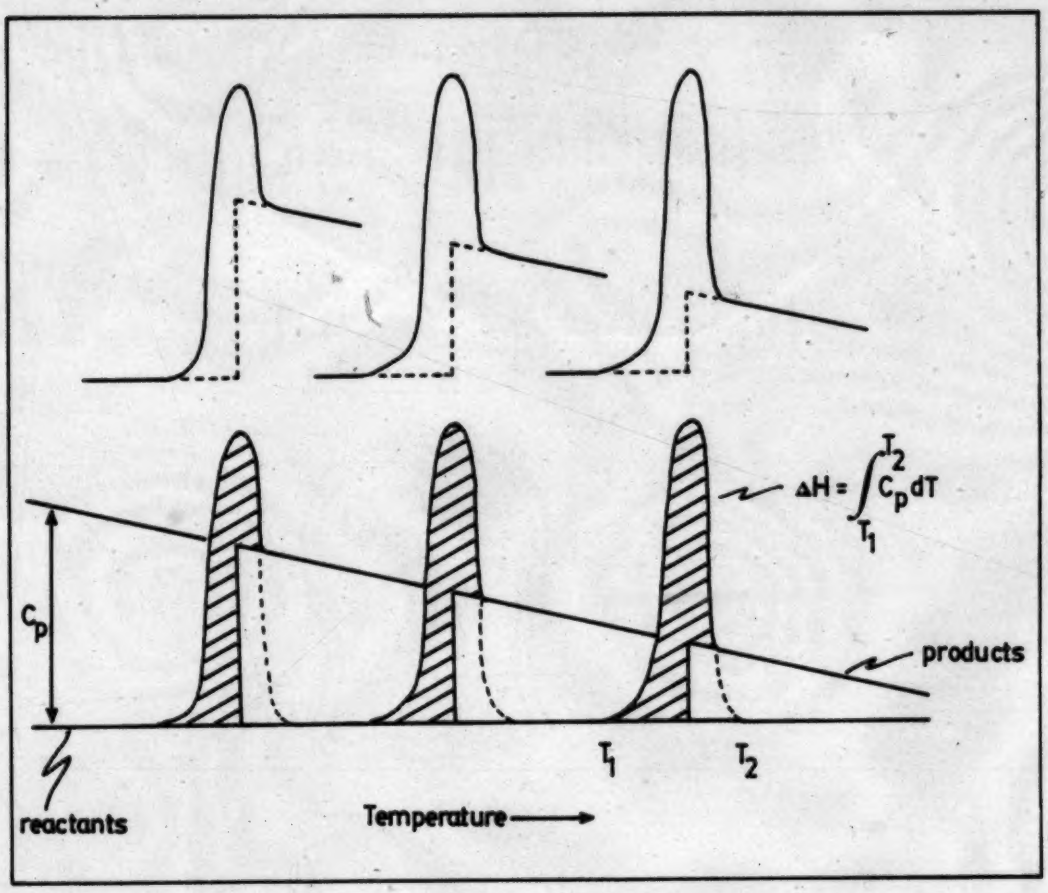


Figure A-2(1) Effect of changes in partial specific heat capacities on apparent melting profiles. See text for discussion. Lower curves indicate underlying events while upper curves show what is observed experimentally.

Ideally one wishes to report the melting enthalpy (hatched area) under standard conditions. If $(\partial \bar{C}_p/\partial T)_p$ differs for single-stranded and double-stranded DNA's then this cannot be determined from a single measurement under non-standard conditions.

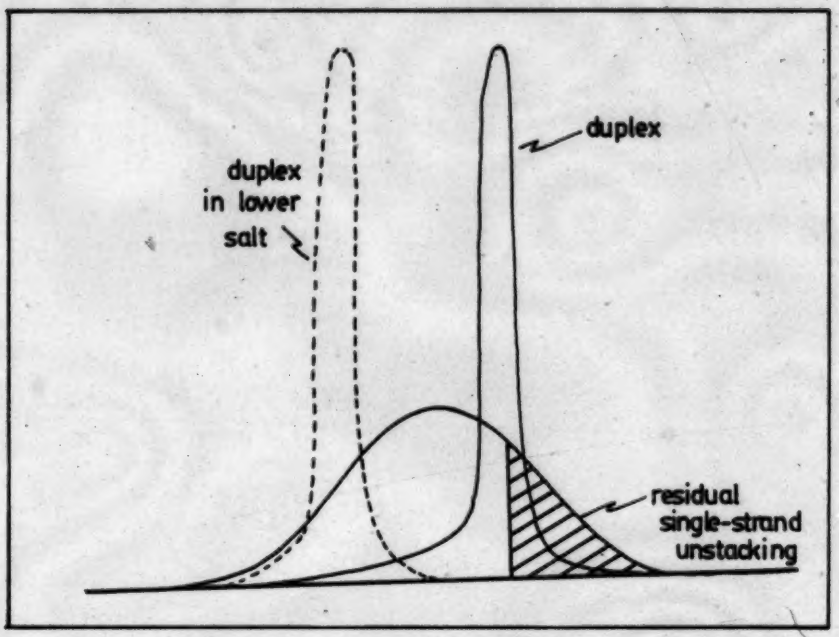


Figure A-2(2) Correction for residual single-strand stacking. Because the double helix will not, in general, melt to free single-strands the residual enthalpy of stacking (hatched area) must be added to the melting enthalpy. See text for discussion.

unstacking transitions are relatively salt independent so the correction varies as the T_m of the duplex is adjusted with salt (Figure A-2 (2)). It is assumed in Chapter VII that residual stacking in the products of DNA melting is small and that the correction need not be applied.

Note the effects described in A and B partially compensate one another in polynucleotides (Filimonov and Privalov, 1978).

BIBLIOGRAPHY

Ackerman, S., Cool, B. and Furth, J.J. Anal. Bioch. 100:174-178 (1979).

Adman, R. and Grossman, L. J. Mol. Biol. 23:417-439 (1967).

Albergo, D.D., et al. Biochemistry 20:1409-1413 (1981).

Alberts, B. and Herrick, G. Methods in Enzymology 21:198-217 (1971).

Ames, B.N., Lee, F.D. and Durston, W.E. Proc. Nat. Acad. Sci. U.S.A. 70:782-786 (1973).

Anderson, C.F., Record, M.T. and Hart, P.A. Biophys. Chem. 7:301-316 (1978).

Arnott, S. and Hukins, D.W.L. Biochem. Biophys. Res. Commun. 47:1504-1509 (1972).

Arnott, S. and Bond, P.J. Science 181:68-69 (1973).

Arnott, S., Chandrasekaran, R. and Leslie, A.G.W. J. Mol. Biol. 106: 738-748 (1976).

Atkins, J.F., et al. Cell 18:1119-1131 (1979).

Bambara, R.A., Vyemura, D., and Choi, T. J. Biol. Chem. 253:413-423 (1978).

Barker, D.L. and Marsh, R.G. Acta Cryst. 17:1581-1587 (1964).

Berman, H.M., et al. Biochim. Biophys. Acta 561:124-131 (1979).

Bleam, M.L., Anderson, C.F. and Record, M.T. Proc. Nat. Acad. Sci. 77:3085-3089 (1980).

Borer, P.N., et al. J. Mol. Biol. 86:843-853 (1974).

Brahms, J., Michelson, A.M. and van Holde, K.E. J. Mol. Biol. 15: 467-488 (1966).

Breslauer, K.J., Sturtevant, J.M. and Tinoco, I. J. Mol. Biol. 99: 549-565 (1975).

Broomhead, J.M. Acta Cryst. 1:324-329 (1948).

Broomhead, J.M. Acta Cryst. 4:92-100 (1951).

Broude, N.E. and Budowsky, R.I. Biochim. Biophys. Acta 254:380-388 (1971).

Bugg, C.E., Thewalt, V.T. and Marsh, R.E. Biochim. Biophys. Res. Commun. 33:436-440 (1968).

Burd, J.F. and Wells, R.D. J. Mol. Biol. 53:435-459 (1970).

Burd, J.F., Larson, J.E. and Wells, R.D. J. Biol. Chem. 250:6002-6007 (1975).

Burgess, R.R. J. Biol. Chem. 244:6160-6167 (1969).

Burgess, R.R. and Jendrisak, J.J. Biochemistry 14:4634-4638 (1975).

Cantor, C.R. and Schimmel, P.R. "Biophysical Chemistry" Parts I, II and III, W.H. Freeman & Co., San Francisco, 1980.

Cassani, G.R. and Bollun, F.J. Biochemistry 8:3928-3936 (1969).

Chamberlin, M.J. and Patterson, D.L. J. Mol. Biol. 12:410-428 (1965).

Ciardelli, T.L., Chakravarty, P.K. and Olsen, R.K. J. Am. Chem. Soc. 100:7684-7690 (1978).

Cohen, J.S., Wooten, J.B. and Chatterjee, C.L. Biochemistry 20:3049-3055 (1981).

Coulter, M.B. "Properties of S Factor", Ph.D. Thesis, University of Alberta (1974).

Crick, F.H.C. J. Mol. Biol. 19:548-555 (1966).

Crothers, D.M. and Zimm, B.H. J. Mol. Biol. 9:1-9 (1964).

Cullis, P.M. and Wolfenden, R. Biochemistry 20:3024-3028 (1981).

DasGupta, C. and Radding, C.M. Nature 295:71-73 (1982).

De Clerq, E., et al. J. Biol. Chem. 250:2521-2531 (1975).

Dickerson, R.E. and Drew, H.R. J. Mol. Biol. 149:761-786 (1981).

DiVerdi, J. and Opella, S.J. J. Mol. Biol. 149:307-312 (1981).

Dodgson, J.B. and Wells, R.D. Biochemistry 16:2367-2374 (1977).

Donohue, J. and Trueblood, K.N. J. Mol. Biol. 2:363-371 (1960).

Donohue, J. Science 165:1091-1096 (1969).

Drake, J.W. "The Molecular Basis of Mutation", Holden-Day Inc. San Francisco 1970.

Drake, J.W. and Baltz, R.H. Ann. Rev. Bioch. 45:11-37 (1976).

Drew, H.R. et al. Nature 286:567-573 (1980).

Drew, H.R. et al. Proc. Nat. Acad. Sci. U.S.A. 78:2179-2183 (1981).

Drew, H.R. and Dickerson, R.E. J. Mol. Biol. 151:535-556 (1981).

Early, T.A. et al. Nucleic Acids Research 5:1955-1970 (1978).

Early, T.A. et al. Biochemistry 20:3756-3763 (1981a).

Early, T.A. et al. Biochemistry 20:3764-3769 (1981b).

Englander, J.J., Kallenbach, N.R. and Englander, S.W. J. Mol. Biol. 63:153-169 (1972).

Ezra et al. Biochemistry 16:1977-1987 (1977).

Feldberg, R.S. and Grossman, L. Biochemistry 15:2402-2408 (1976).

Feldberg, R.S. Nucleic Acids Research 8:1133-1144 (1980).

Feldberg, R.S. in "Chromosome Damage and Repair" (ed. Seeberg, E.) Plenum Press, New York (in press).

Filimonov, V.V. and Privalov, P.L. J. Mol. Biol. 122:465-470 (1978).

Finch, J.T. and Klug, A. J. Mol. Biol. 46:597-598 (1969).

Fink, T.R. and Crothers, D.M. J. Mol. Biol. 66:1-12 (1972).

Fresco, J.R. and Alberts, B.M. Proc. Nat. Acad. Sci. U.S.A. 46:311-321 (1960).

Frier, S.M. et al. Biochemistry 20:1419-1426 (1981).

Fuller, W., Wilkins, M.H.F., Wilson, H.R. and Hamilton, L.D. J. Mol. Biol. 12:60-76 (1965).

Garrett, R.A., Douthwaite, S. and Noller, H.F. Trends Biochem. Sci. 6:137-139 (1981).

Geerdes, H. and Hilbers, C.W. FEBS Lett. 107:125-128 (1979).

Gellert, M., Lipsett, M.N. and Davies, D.R. Proc. Nat. Acad. Sci. U.S.A. 48:2013-2018 (1962).

Gerdil, R. Acta Cryst. 14:333-344 (1961).

Greene, J.J. et al. Biochemistry 17:4214-4220 (1978).

Ghangas, G.S. et al. Biochim. Biophys. Res. Commun. 54:998-1007 (1973).

Gorenstein, D.G. Ann. Rev. Biophys. Bioeng. 10:355-386 (1981).

Gorenstein, D.G. in NMR Spectroscopy in Molecular Biology, (ed. Pullman, B.) D. Reidel Publishing Co., Dordrecht (Holland) (1978).

Gorenstein, D.G. et al. Biochemistry 20:2141-2150 (1981).

Gray, D.M. and Ratliff, R.L. Biopolymers 16:1331-1342 (1977).

Gray, D.M. et al. Nucleic Acids Research 8:3695-3707 (1980).

Gray, H.B. et al. Nucleic Acids Research 3:1459-1492 (1975).

Haasnoot, C.A.G. et al. Nucleic Acids Research 8:169-181 (1980).

Hamilton, W.C. and Ibers, J.A. "Hydrogen-bonding in solids", Benjamin, New York, 1968.

Hanlon, S. Biochim. Biophys. Res. Commun. 23:861-867 (1966).

Hartel, A.J. et al. Nucleic Acids Research 9:1405-1423 (1981).

Harwood, S.J. and Wells, R.D. J. Biol. Chem. 245:5625-5634 (1970).

Haschemeyer, A.E.V. and Sobell, H.M. Proc. Nat. Acad. Sci. U.S.A. 50:872-877 (1963).

Helfgott, D.C. and Kallenbach, N.R. Nucleic Acids Research 7:1011-1017 (1979).

Hirose, S. et al. J. Mol. Biol. 77:501-517 (1973).

Hogan, M.E. and Jardetsky, O. Biochemistry 19:3460-3468 (1980).

Holbrook, S.R. et al. J. Mol. Biol. 123:631-660 (1978).

Holcomb, D.N. and Tinoco, I. Biopolymers 3:121-133 (1965).

Hoogsteen, K. Acta Cryst. 12:822-823 (1959).

Hori, H. and Osawa, S. Proc. Nat. Acad. Sci. U.S.A. 76:381-385 (1979).

Howard, F.B. and Miles, H.T. Biochemistry 16:4647-4650 (1977).

Ifft, J.B., Voet, D.H. and Vinograd, J. J. Phys. Chem. 65:1138-(1961).

Inman, R.B. J. Mol. Biol. 9:624-637 (1964).

Inman, R.B. and Baldwin, R.L. J. Mol. Biol. 8:452-469 (1964).

Jack, A. et al. J. Mol. Biol. 108:619-649 (1976).

Jager, J.J. in "Introduction to Research in Ultraviolet Photobiology" Prentice-Hall Inc., London (1967).

Jencks, W.P. "Catalysis in Chemistry and Enzymology", McGraw-Hill, New York, 1967.

Johnson, D. and Morgan, A.R. Proc. Nat. Acad. Sci. U.S.A. 75:1637-1641 (1978).

Jovin, T.M. et al. J. Biol. Chem. 244:2996-3008 (1969).

Jovin, T.M. Ann. Rev. Bioch. 45:889-920 (1976).

Karran, P. and Lindahl, T. Biochemistry 19:6005-6011 (1980).

Katz, L., Tomita, K. and Rich, A. J. Mol. Biol. 13:340-350 (1965).

Katz, L. and Penman, S. J. Mol. Biol. 15:220-231 (1966).

Kelly, T.J. and Thomas, C.A. J. Mol. Biol. 44:459-475 (1969).

Kimhi, Y. and Littauer, U.Z. J. Biol. Chem. 243:231-240 (1968).

Klug, A. et al. J. Mol. Biol. 131:669-680 (1979).

Kornberg, A. "DNA Replication" W.H. Freeman & Co., San Francisco, 2nd Edition 1980.

Kornberg, A. et al. Proc. Nat. Acad. Sci. U.S.A. 51:315-323 (1964).

Krakauer, H. and Sturtevant, J.M. Biopolymers 6:491-512 (1968).

Kubinski, H., Opara-Kubinska, Z. and Szybalski, W. J. Mol. Biol. 20:313-329 (1966).

Kunitz, M. J. Gen. Physiol. 33:349- (1950).

Kyogoku, Y., Lord, R.C. and Rich, A. Science 154:518-520 (1966).

Kyogoku, Y., Lord, R.C. and Rich, A. Proc. Nat. Acad. Sci. U.S.A. 57:250-257 (1967).

Kyogoku, Y. et al. Biochim. Biophys. Acta 179:10-17 (1969).

Lagerkvist, U. Cell 23:305-306 (1981).

Langridge, R. et al. J. Mol. Biol. 2:19-37 (1960a).

Langridge, R. et al. J. Mol. Biol. 2:38-64 (1960b).

Lee, C.-H. and Tinoco, I. Nature 274:609-610 (1978).

Lee, C. et al. Biochemistry 15:3627-3639 (1976).

Lee, J.S. and Waring, M.J. Biochem. J. 173:129-144 (1978).

Lee, J.S., Johnson, D. and Morgan, A.R. Nucleic Acids Research 6:3073-3091 (1979).

Lee, J.S., Evans, D.H. and Morgan, A.R. Nucleic Acids Research 8:4305-4320 (1980).

Lee, J.S. et al. Nucleic Acids Research 9:1707-1721 (1981).

Legerski, R.J., Gray, H.B. and Robberson, D.L. J. Biol. Chem. 252: 8740-8746 (1977).

Lehman, I.R. and Nussbaum, A.C. J. Biol. Chem. 239:2628-2636 (1964).

Leng, M. and Felsenfeld, G. J. Mol. Biol. 15:455-466 (1966).

Leslie, A.G.W. et al. J. Mol. Biol. 143:49-72 (1980).

Levitt, M. and Warshel, A. J. Am. Chem. Soc. 100:2607-2613 (1978).

Levitt, M. Proc. Nat. Acad. Sci. U.S.A. 75:640-644 (1978).

Linn, S. and Lehman, I.R. J. Biol. Chem. 240:1287-1293 (1965a).

Linn, S. and Lehman, I.R. J. Biol. Chem. 240:1294-1304 (1965b).

Lippke, J.A. et al. Proc. Nat. Acad. Sci. U.S.A. 78:3388-3392 (1981).

Lipsett, M.N. J. Biol. Chem. 239:1256-1260 (1964).

Lomant, A.J. and Fresco, J.R. Prog. Nucleic Acids Research and Mol. Biol. 12:1-27 (1972).

Lomant, A.J. and Fresco, J.R. J. Mol. Biol. 77:345-354 (1973).

Lomant, A.J. and Fresco, J.R. Prog. Nucleic Acids Research and Mol. Biol. 15:185-218 (1975).

Losick, R. and Chamberlin, M. (eds.) "RNA Polymerase" Cold Spring Harbour Laboratory, Cold Spring Harbour, New York, (1976).

Mabrey, S. and Sturtevant, J.M. Methods in Membrane Biology 9:237-274 (1978).

Manning, G.S. Biopolymers 11:937-949 (1972a).

Manning, G.S. Biopolymers 11:951-955 (1972b).

Manning, G.S. Quart. Rev. Biophys. 11:179-246 (1978).

Mandal, C., Kallenbach, N.R. and Englander, S.W. J. Mol. Biol. 135: 391-411 (1979).

Martin, F.H., Uhlenbeck, O.C. and Doty, P. J. Mol. Biol. 57:201-215 (1971).

Marvin, D.A., Spencer, M., Wilkins, M.H.F., Hamilton, L.D. J. Mol. Biol. 3:547-565 (1961).

Masamune, Y., Fleischman, R.A. and Richardson, C.C. J. Biol. Chem. 246:2680-2691 (1971).

Mathews, F.S. and Rich, A. J. Mol. Biol. 8:89-95 (1964).

McCutchan, T.F. et al. Nucleic Acids Research 2:853-864 (1975).

McEntee, K., Weinstock, G. and Lehman, I.R. Proc. Nat. Acad. Sci. U.S.A. 77:857-861 (1980).

McGhee, J.D. Biopolymers 15:1345-1375 (1976).

McGhee, J.D. and von Hippel, P.H. J. Mol. Biol. 86:469-489 (1974).

McGhee, J.D. and von Hippel, P.H. J. Mol. Biol. 103:679 (1976).

Metzler, D.E. "Biochemistry" Academic Press, New York, 1977.

Mickel, S., Arena, V. and Bauer, W. Nucleic Acids Research 4:1465-1481 (1977).

Miles, H.T. and Frazier, J. Biochim. Biophys. Acta. 79:216-220 (1964).

Milman, G., Chamberlin, M. and Langridge, R. Proc. Nat. Acad. Sci. U.S.A. 57:1804-1810 (1967).

Morgan, A.R. and Wells, R.D. J. Mol. Biol. 37:63-80 (1968).

Morgan, A.R. J. Mol. Biol. 52:441-466 (1970).

Morgan, A.R. et al. Biochemistry 13:1596-1603 (1974).

Morgan, A.R. et al. Nucleic Acids Research 7:547-570 (1979a).

Morgan, A.R. et al. Nucleic Acids Research 7:571-594 (1979b).

Murray, N.L. "Three-Stranded Polynucleotide Structures", M.Sc. Thesis, University of Alberta, 1972.

Nagel, G.M. and Hanlon, S. Biochemistry 11:816-823 (1972a).

Nagel, G.M. and Hanlon, S. Biochemistry 11:823-830 (1972b).

Neumann, E. and Ackermann, T. J. Phys. Chem. 71:2377-2379 (1967).

Nishimura, S., Jones, D. and Khorana, H.G. J. Mol. Biol. 13:302-324 (1965).

O'Brian, E.J. J. Mol. Biol. 7:107-110 (1963).

Olson, K., Luk, D. and Harvey, C.L. Biochim. Biophys. Acta 277:269-275 (1972).

Paetkau, V., et al. J. Mol. Biol. 71:293-306 (1972).

Panayotatos, N. and Wells, R.D. Nature 289:466-470 (1981).

Parry, G.S. Acta Cryst. 7:313-320 (1954).

Patel, D.J. and Canuel, L.L. Eur. J. Biochem. 96:267-276 (1979).

Pauling, L. and Corey, R.B. Arch. Biochem. Biophys. 65:164-181 (1956).

Pohl, F.M. and Jovin, T.M. J. Mol. Biol. 67:375-396 (1972).

Pribnow, D., et al. J. Mol. Biol. 149:337-376 (1981).

Privalov, P.L., et al. Biopolymers 8:559-571 (1969).

Privalov, P.L. and Filimonov, V.V. J. Mol. Biol. 122:447-464 (1978).

Quigley, G.J. and Rich, A. Science 194:796-806 (1976).

Rabin, E.Z. and Fraser, M.J. Can. J. Biochem. 48:389-392 (1970).

Radding, C.M., Josse, J. and Korberg, A. J. Biol. Chem. 237:2869-2876 (1962).

Randerath, K. and Randerath, E. Methods in Enzymology 12:323-347 (1967).

Record, T. J. Mol. Biol. 107:145-158 (1976).

Record, T., Anderson, C.F. and Lohman, T.M. Rev. Biophys. 11:103-178 (1978).

Record, M.T., et al. Ann. Rev. Biochem. 50:997-1024 (1981).

Reid, B.R. Ann. Rev. Biochem. 50:969-996 (1981).

Rhodes, D. and Klub, A. Nature 286:573-578 (1980).

Rich, A., et al. J. Mol. Biol. 3:71-86 (1961).

Richardson, C.C., Lehman, I.R. and Kornberg, A. J. Biol. Chem. 239:

251-258 (1964).

Richardson, C.C., et al. J. Biol. Chem. 239:222-232 (1964).

Riley, M., Maling, B. and Chamberlin, M.J. J. Mol. Biol. 20:359-389 (1966).

Ringer, D.P., Burchett, S. and Kizer, D.E. Biochemistry 17:4818-4825 (1978).

Romaniuk, P.J., et al. Biochemistry 18:5109-5116 (1979).

Rubin, J., Brennan, T. and Sunderalingam, M. Science 174:1020-1021 (1971).

Sakore, R.D., Travale, S.S. and Sobell, H.M. J. Mol. Biol. 43:361-374 (1969).

Salser, W. Cold Spring Harb. Symp. Quant. Biol. 42:985-1002 (1977).

Schachman, H.K., et al. J. Biol. Chem. 235:3242-3249 (1960).

Seeman, N.C., et al. J. Mol. Biol. 104:109-144 (1976).

Selsing, E., Wells, R.D., Early, T.A. and Kearns, D.R. Nature 275: 249-250 (1978).

Setlow, R.B. and Polland, E.C. "Molecular Biophysics", Addison Wesley, Reading, Mass., 1962.

Setlow, R.B. and Carrier, W.L. J. Mol. Biol. 17:237-254 (1966).

Setlow, P., Brutlag, D. and Kornberg, A. J. Biol. Chem. 247:224-231 (1972).

Stevens, C.L. and Rosenfeld, A. Biochemistry 5:2714-2715 (1966).

Streisinger, G., et al. Cold Spring Harb. Symp. Quant. Biol. 31: 77-84 (1966).

Shapiro, J.T., Stannard, B.S. and Felsenfeld, G. Biochemistry 8:3233-3241 (1969).

Shindo, H., et al. Biochemistry 19:518-526 (1980).

Sigal, N. and Alberts, B. J. Mol. Biol. 71:789-793 (1972).

Skoog, D.A. and West, D.M. "Fundamentals of Analytical Chemistry", Holt, Rinehart and Winston, 1976, (3rd edition), New York.

Smith, D.A., et al. J. Biol. Chem. 242:590-595 (1967).

Solie, T.N. and Schellman, J.A. J. Mol. Biol. 33:61-77 (1968).

Springate, C.F. and Loeb, L.A. J. Mol. Biol. 97:577-591 (1975).

Studier, F.W. J. Mol. Biol. 11:373-390 (1965).

Sundaralingam, M. "Structure and Conformation of Nucleic Acids and Protein - Nucleic Acid Interactions" (eds. M. Sundaralingam and S.T. Roo), University Park Press, Baltimore, 1975.

Sussman, J.L., et al. J. Mol. Biol. 123:607-630 (1978).

Szybalski, W. Methods in Enzymology 12:330-360 (1968).

Tavale, S.S., Sakore, T.D. and Sobell, H.M. J. Mol. Biol. 43:375-384 (1969).

Tazawa, I., Tazawa, S. and Ts'o, P.O.P. J. Mol. Biol. 66:115-130 (1972).

Thiele, D. and Guschlbauer, W. Biophysik 9:261-277 (1973).

Thomas, M. and Davis, R.W. J. Mol. Biol. 91:315-328 (1975).

Tinoco I., et al. Nature (New Biology) 246:40-41 (1973).

Topal, M.D. and Fresco, J.R. Nature 263:285-289 (1976a).

Topal, M.D. and Fresco, J.R. Nature 263:289-293 (1976b).

Topal, M.D. and Fresco, J.R. Biochemistry 19:5531-5537 (1980).

Ts'o, P.O.P., Melvin, I.S. and Olson, A.C. J. Am. Chem. Soc. 85: 1289-1296 (1963).

Ts'o, P.O.P., et al. Biochemistry 5:4153-4170 (1966).

Varghese, A.J. Biochemistry 9:4781-4787 (1970).

Varghese, A.J. Biochemistry 10:2194-2199 (1971).

Vetterl, V. and Guschlbauer, W. Arch. Biochem. Biophys. 148:130-140 (1972).

Viswamitra, M.A., et al. Nature 273:687-688 (1978).

Viswamitra, M.A., et al. Nature 289:817-819 (1981).

Voet, D. and Rich, A. Prog. Nucleic Acids Research and Mol. Biol. 10:183-265 (1970).

Vogt, V.M. Eur. J. Biochem. 33:192-200 (1973).

Wang, A.C. and Kallenbach, N.R. J. Mol. Biol. 62:591-611 (1971).

Wang, A.H-J., et al. Nature 282:680-686 (1979).

Wang, J.C. Proc. Nat. Acad. Sci. U.S.A. 76:200-203 (1979).

Waring, M.J. Biochem. J. 143:483-486 (1974).

Wartell, R.M., Larson, J.E. and Wells, R.D. J. Biol. Chem. 249:6719-6731 (1974).

Watson, D.G., et al. Acta Cryst. 19:111-124 (1965).

Watson, J.D. and Crick, F.H.C. Nature 171:737 (1953).

Weith, H.L., Wiebers, J.L. and Gilham, P.T. Biochemistry 9:4396-4401 (1970).

Wells, R.D., et al. J. Mol. Biol. 27:237-263 (1967).

Wells, R.D., et al. J. Mol. Biol. 54:465-497 (1970).

Wells, R.D., et al. CRC Critical Reviews in Biochemistry 4:305-340 (1977).

Wells, R.D., et al. J. Biol. Chem. 254:5535-5541 (1979).

Westhof, E., Rao, S.T. and Sundaralingam, M. J. Mol. Biol. 142:331-361 (1980).

Wing, R.M., et al. Nature 287:755-758 (1980).

Wolfenden, R.V. J. Mol. Biol. 40:307-310 (1969).

Yonuschot, G. and Mushrush, G.W. Biochemistry 14:1677-1681 (1975).

Young, P.R. and Kallenback, N.R. J. Mol. Biol. 126:467-479 (1978).

Zimmerman, S.B., Cohen, G.H. and Davies, D.R. J. Mol. Biol. 92:181-192 (1975).

Zimmerman, S.B. and Pheiffer, B.H. Proc. Nat. Acad. Sci. U.S.A. 78: 78-82 (1981).